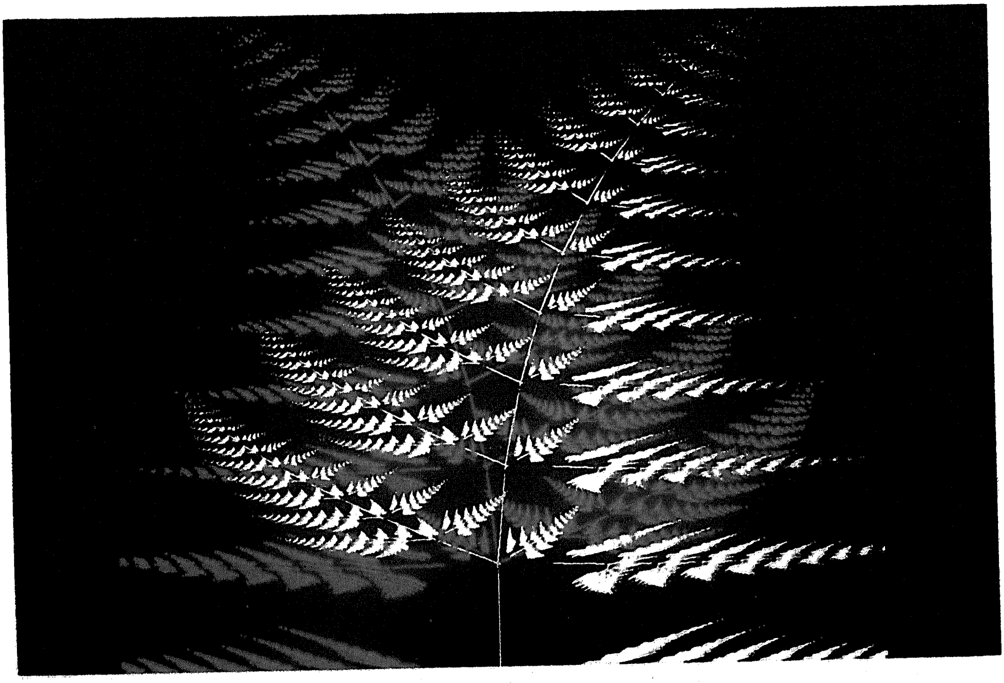
Resonance

May 1996 Volume 1 Number 5

journal of science education



Chaos Modelling with Computers ❖ River Piracy: Saraswati that Disappeared ❖ Evidence for Bird Mafia ❖ Java Internet Programming Language ❖ Curves vs Surfaces

Resonance is a monthly journal of science education published by Indian Academy of Sciences, Bangalore, India.

Editorial Office: Indian Academy of Sciences, C V Raman Avenue, PB No. 8005, Bangalore 560 080, India.
Tel:+91 (80) 3342310 / 3342456, Fax:+91 (80) 3346094, email:resonanc@ias.ernet.in

Editorial

N Mukunda, Chief Editor

There is an age-old legend that at the confluence of the rivers Ganga and Yamuna at Prayag (Allahabad), a third unseen river Saraswati joins the other two. What, where and who is this Saraswati? In this issue, Khadg Singh Valdiya describes in delightful fashion the evidence to show that not too long ago there was indeed a major river by that name, and explains how she disappeared. You will also be amused to discover the connotation of the name 'Yamuna' for a river, as Valdiya explains by looking at examples from different parts of India.

We have, partly by design, evolved a pattern of telling you something interesting in each issue about some major figure in science or mathematics, by having a portrait on the back cover, accompanied elsewhere by a brief account of the person. This time we feature the mathematician John von Neumann. But to call him a mathematician is to unfairly limit his genius and his accomplishments! His class-fellow and later world-renowned theoretical physicist Eugene Wigner once described him as "the sharpest intellect whom I have known". In turn, it was von Neumann who suggested to Wigner that the mathematical discipline of group theory could be applied with great profit to the then new subject of quantum mechanics. And this Wigner did with outstanding success. You will find some aspects of this reflected in Vishwambhar Pati's review of Shlomo Sternberg's book "Group Theory and Physics".

Several readers have written recently suggesting that we reduce the amount of material presented in each issue. In response (and for other reasons too!), we have made the issues starting with April 1996 somewhat slimmer than earlier ones. So now you might say - 'Yond ***Resonance*** has a lean and healthy look'!

In response to suggestions from several readers (and for other reasons too!), we have made the issues starting with April 1996 somewhat slimmer than earlier ones. So now you might say - 'Yond *Resonance* has a lean and healthy look'!

The First Electronic Computer

1996 is the fiftieth anniversary of the birth of the first electronic computer. On February 14, 1946 the Electronic Numerical Integrator and Computer (ENIAC) was formally switched on at the Moore School of Electrical Engineering at the University of Pennsylvania, U.S.A. ENIAC, designed by a team headed by John W Mauchly and J Persper Eckert Jr. used 18000 vacuum tubes, weighed 30 tonnes, occupied a 10mx15m room and took 3 years to build. Its main goal was to calculate the trajectories of missiles. Precedence in designing the first electronic computer is claimed by John Atanasoff and Clifford E Berry, who in 1941 had designed and partially completed an electronic calculator which used 300 vacuum tubes to add and subtract. John Mauchly had met Atanasoff and Berry at Iowa State College in 1941 and did gain from their discussions on building ENIAC (Atanasoff's machine was not completed due to the exigencies of war). ENIAC thus became the first large electronic computer successfully used for solving important problems.

ENIAC was programmed by plugging wires on a large 1 sq.m plug board which interconnected various arithmetic circuits. Each program required a different plug board to be wired and this was a tedious job.

John von Neumann became involved with the ENIAC team in August 1944 by which time the difficulties of programming ENIAC were quite evident. Von Neumann, Eckert and Mauchly cooperated in initiating the design of a successor to ENIAC called EDVAC (Electronic Discrete Variable Automatic Computer). In June 1945, von Neumann wrote a report titled 'First draft report on the EDVAC' in which he examined the problem of computer design logically and identified design principles which went beyond the electronic hardware problems of the day. He proposed building a *stored program computer* in which instructions and data would be stored in the same storage unit that he called *memory,* invoking neurological terminology. The idea of storing data and instructions indistinguishably in the same memory was a master stroke. This allowed one to repetitively execute a sequence of instructions with different data in each repetition leading to concise programs. By treating instructions as data, they could be altered based on previous computations thereby 'adaptively' altering a program. Over the last 50 years, advances in technology have made computers smaller, cheaper and extremely fast. At a fundamental level, however, they are all stored program computers whose architecture was originally proposed by von Neumann.

Suggested Reading

N Stern. From ENIAC to UNIVAC. Digital Press. Bedford. MA. USA. 1981.

M R Williams. A History of Computing Technology. Prentice Hall Inc. Englewood Cliffs, U.S.A., 1985.

V Rajaraman

Science Smiles

R K Laxman

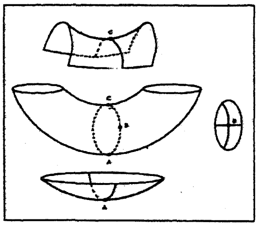
There, I've succeeded in quickening the evolutionary process!
Now he will conduct experiments on us!

Resonance journal of science education

May 1996 Volume 1 Number 5

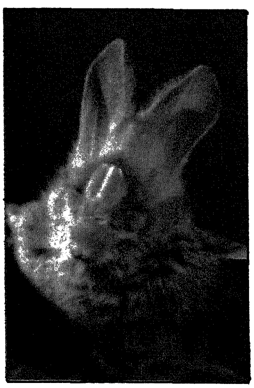
SERIES ARTICLES

6

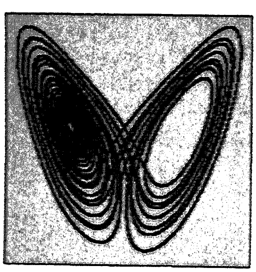


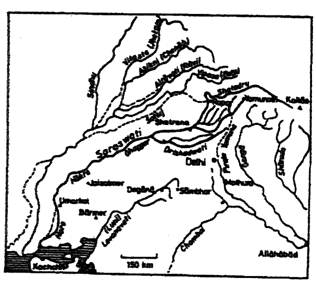
GENERAL ARTICLES

40

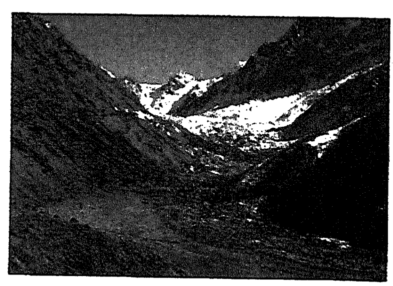


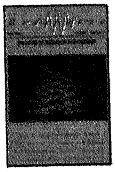
29





19

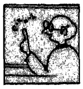


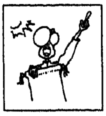


Front Cover
The fern on the cover is a computer generated fractal. It was programmed by C E Prakash of the Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, on a Silicon Graphics workstation. The method of generation is described by M F Barnsley in his book 'Fractals Everywhere' (Academic Press, New York, 1988).

Back Cover
John (Janos) von Neumann (1903 - 1957)
(illustration by V Pati)

DEPARTMENTS





Geometry

4. Curves vs Surfaces

Kapil H Paranjape

After spending about a decade at the School of Mathematics, TIFR, Bombay, Kapil H Paranjape is currently with Indian Statistical Institute, Bangalore.

In the first three parts the author covered what can roughly be called 'classical' geometry. These are the aspects of geometry that all science and engineering graduates have at some time studied. Unfortunately the 'modern' material you will now encounter, which is a beautiful combination of ideas from algebra and analysis, has not entirely made it to the common curriculum. So arm yourselves with a paper and pencil[0] and begin your excursion.

[0] The readers are encouraged to solve or at least attempt all exercises given since (to quote a famous mathematician) 'mathematics is not banana eating.'

Is a Curve Curved?

Curve: A one-dimensional locus of points[1].

[1] This is a fancy way of saying 'that which is traced by the tip of a moving finger'—which having writ moves on.

Now that we have accepted coordinate geometry as the basis of our study, we need to re-examine earlier notions like lines and planes in this context. In some sense, the more natural notion is that of a curve rather than a line. Huygens, Leibnitz and Newton (independently) formulated the notion of curvature of a curve. (This was developed by Serret-Frenet into a multiplicity of invariants for curves in higher dimensions. We will concentrate on the curvature defined by Huygens et al). A line then becomes a curve of curvature zero.

It is always confusing to look at the dictionary meaning of a word when it is also a mathematical term. The word 'line' in English corresponds to the mathematical term 'curve', whereas the English word 'curve' is used in the sense of a line which is *not* straight. A 'straight line' (in English) is what one would call a 'line' in geometry. As we shall see it is not easy to distinguish a curve from a straight line intrinsically—thus we use the mathematical term curve to denote a one dimensional locus which *may or may not be* curved!

Since we will only be dealing with non-singular[2] loci we use the analytic definition that near each point $p_0 = (x_0, y_0, z_0)$ of the curve it is given in parametric form as $p(t) = (x(t), y(t), z(t))$ where these coordinates are given by regular functions of t near 0. Then the non-singularity of the locus translates to the non-vanishing of the velocity vector $(\dot{x}(t), \dot{y}(t), \dot{z}(t))$.

[2] Mathematicians are prone to describing things which are easy to study by nice names like 'regular' while those which are difficult are given epithets like 'singular'.

In order to understand how such a locus is curved, we travel along it at constant speed and check to see whether we experience any acceleration. This is in keeping with Newton's law that a body moves along a (straight) line at constant speed unless it is subjected to acceleration; thus Newton's law will turn out to be a geometric *definition* of a line. In symbols, let $p(t) = (x(t), y(t), z(t))$ be a parametrisation of the curve so that we have speed equal to a constant, i.e.

$$(\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2) = \text{constant independent of } t.$$

(Exercise: Those who know their calculus should be able to show that such a parametrisation always exists). The acceleration is then a vector perpendicular to the velocity

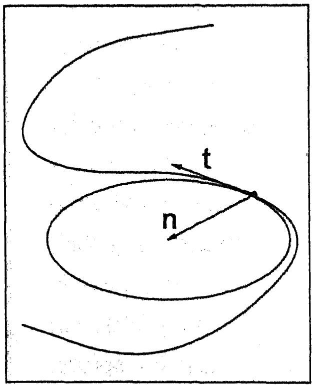
$$\dot{x}(t)\ddot{x}(t) + \dot{y}(t)\ddot{y}(t) + \dot{z}(t)\ddot{z}(t) = 0.$$

(Exercise: Why does this equation hold?) The magnitude $k(t) = (\ddot{x}(t)^2 + \ddot{y}(t)^2 + \ddot{z}(t)^2)^{1/2}$ of the acceleration is called the curvature of the curve at the point $p(t)$. If one works this out for the circle

$$(x(t), y(t), z(t)) = (r\cos(t), r\sin(t), 0)$$

then we obtain $k(t) = 1/r$ (Exercise: Check this). This corroborates our intuition that a circle with a large radius is nearly a straight line, i.e. has curvature close to zero.

Figure 1 ***The osculating circle. (t: tangent vector; n: normal vector.)***



A more geometric approach is as follows. Just as the tangent line at a point is given by the linear equation that approximates the curve upto terms of order 2 or more, the *osculating circle* is the circle that approximates the curve upto terms of order 3 or more (see *Figure 1*). The curvature of the curve at the point is defined to be the inverse of the radius of the osculating circle (here we

allow a circle to *become* a line when its radius is infinite). It is not hard to show that the analytic definition coincides with the geometric one. This points the way to the higher order invariants of Serret-Frenet which are obtained by examining curves of higher degree which approximate the given one to even higher orders.

One problem is that this curvature is not *intrinsic*. An object that travels along the curve *without* interacting with the surroundings[3] will not observe the acceleration since the latter is perpendicular to the curve. One way of seeing this is to think of a signal moving in a TV cable or optical telephone fibre cable—the coiling or straightening of the cable has no effect on the signal. It is reasonably clear from this viewpoint that there is *no* intrinsic notion of curvature for a curve; curvature for a curve is determined by the manner in which the curve is embedded in (sits in) space—the beauty of a curve is truly in the eye of the beholder! To confront intrinsic curvature (and beauty) one must study higher dimensional loci.

[3]Like the arrow of Arjun that travelled to its target looking neither left nor right.

Curvature is Superficial!

Surface: A two-dimensional locus of points or a one-dimensional locus of curves.

Analogous to the case of curves, the analytic definition of a surface is given by a vector valued function $p(u, v) = (x(u, v), y(u, v), z(u, v))$ such that the two vectors

$$\frac{\partial p(u,v)}{\partial u} = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) \text{ and } \frac{\partial p(u,v)}{\partial v} = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right)$$

are linearly independent (i.e. one is not a multiple of the other). The study of curvature of such loci was initiated by Euler and Meusnier and carried to its "remarkable" conclusion (Theorema Egregium) by Gauss.

There is *no* intrinsic notion of curvature for a curve; curvature for a curve is determined by the manner in which the curve is embedded in (sits in) space.

We have already introduced the notion of curvature for any curve

on the surface. We can expect that the curvature of a surface would be determined by understanding the curvature of the curves on it. In particular, as Euler did, one may consider a curve on the surface which appears 'straight' on the surface, i.e. so that the acceleration experienced while travelling along this curve is perpendicular to the surface. These curves describe the distance minimising paths between points on the surface (at least in a small region). One may think about a person walking along a 'straight line' on the surface of the earth and travelling at constant speed; no acceleration is required (assuming that there is no friction) but of course there is an unnoticed acceleration due to gravity which holds the person to the surface of the earth. Because of this description such curves are called *geodesics*. Clearly the curvature of geodesics would be directly related to the curvature of the surface; in fact Meusnier showed that the curvature of other curves on the surface can be easily determined once we know the curvature of geodesics.

If we have two surfaces and an identification between the two so that distances are preserved, then the curvatures must also be the same. This is why there can never be a perfect map of the surface of the earth or a perfect astronomical chart.

Each geodesic emanating from a point p is determined by its initial velocity, which is a linear combination of the basic tangent vectors,

$$\mathbf{t} = a\frac{\partial p}{\partial u} + b\frac{\partial p}{\partial v} = a\left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right) + b\left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right).$$

Fix a vector $\mathbf{n}$ of unit length that is perpendicular to all these vectors. The acceleration experienced on a geodesic along $\mathbf{t}$ is then $k(a,b)\mathbf{n}$ for some scalar function $k(a,b)$ of the two parameters determining this vector. The result proved by Euler was that $k(a,b) = L a^2 + 2Mab + Nb^2$. Normalising $k(a,b)$ by the square of the length of the vector $\mathbf{t}$ we obtain a function of the direction alone. Let k_1 be the maximum normalised value of $k(a,b)$ so obtained, say along $\mathbf{t}_1$. Let $\mathbf{t}_2$ be orthogonal to $\mathbf{t}_1$ so that $\mathbf{t}_1 \times \mathbf{t}_2 = \mathbf{n}$ and let k_2 be the normalised value of $k(a,b)$ in this direction; it can be shown that this is the minimum normalised value. The numbers k_1 and k_2 were called the principal curvatures by Euler.

Gauss (who was perhaps inspired by astronomy or by the

cartographic surveys he was carrying out for the ruler of Germany) gave a new interpretation to Euler's theory. First consider the length of the vector **t** considered as a function $g(a,b)$ of the two parameters. This has the form $g(a,b) = E(u,v)\,a^2 + 2\,F(u,v)\,ab + G(u,v)b^2$ and determines distances along paths on the surface—hence it is intrinsic. The second fundamental form is $k(a,b)$ introduced above. The curvature $K = k_1\, k_2$ (now called Gaussian curvature in his honour) is then the ratio of the discriminants of these two forms

$$K = \frac{LN - M^2}{EG - F^2}$$

Even an ant crawling along a surface (or a cartographer in the days before aerial travel) can determine the curvature. Hence the beauty of a surface is skin deep and yet is naturally associated with it!

The key results are the following:

- The Gaussian curvature can be expressed in terms of E, F, G and their partial derivatives alone. Hence it is an *intrinsic* invariant.
- The integral of the Gaussian curvature on a triangle whose sides are geodesics is [(the sum of the angles of the triangle}) $-\pi$] x (the area of the triangle).

  Hence again it is *intrinsic*.

Exercise: What are k_1 and k_2 for a cylinder of radius r? As a more challenging exercise the reader is invited to check that the points A, B, C on the cycle tube pictured in *Figure 2* exhibit positive, zero and negative Gaussian curvature respectively.

What is new and remarkable here is that the (Gaussian) curvature is an *intrinsic* invariant. In other words, if we have two surfaces and an identification between the two so that distances are preserved, then the curvatures must also be the same. This is why there can never be a perfect map of the surface of the earth or a perfect astronomical chart—a curved surface cannot be identified with a flat one in a manner that preserves distances. The second result of Gauss allows one to compute the average value of curvature in a small region of the surface. Thus even an ant

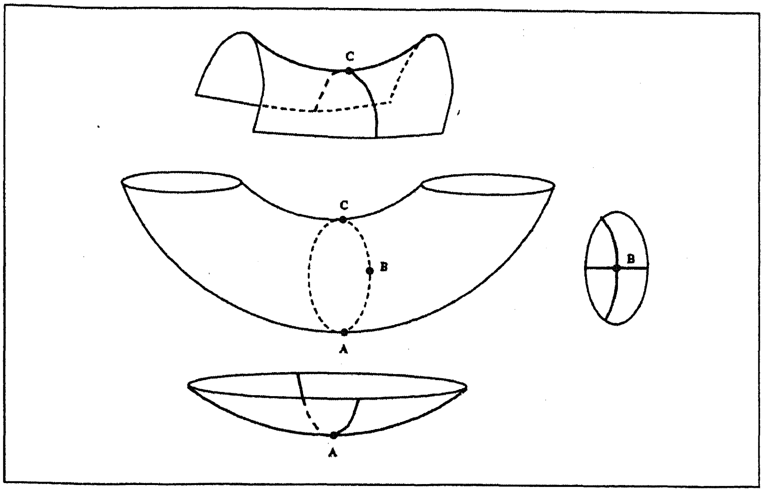


Figure 2 Curvature on a piece of cycle tube. (The thick lines denote the geodesics of extremal curvature.)

crawling along a surface (or a cartographer in the days before aerial travel) can determine the curvature. Hence the beauty of a surface is skin deep and yet is naturally associated with it!

Summary

While curvature is a property which seems to be emanating out of curves, these are in fact not intrinsically curved—the curvature of a curve is entirely the result of how the curve lies in space. However, the intrinsic curvature of a surface (the Gaussian curvature) is a quantity that can be measured without reference to the 'outside'. This curvature relates to the sum of angles of a triangle which is closely related to the parallel postulate.

In spite of all the positive features of Gaussian curvature it has one major drawback which is that its definition is not intrinsic since it depends on the extrinsic curvatures of curves contained in the surface. However, since it is an intrinsic invariant one should expect an intrinsic definition. Further, it is not clear what the correct analogue in higher dimensions should be (in particular in the all important dimension 3). Historically, it was philosophically unacceptable to compute invariants for space by postulating a higher dimensional universe in which it is contained. It was Riemann under the insistence of Gauss who produced the definitive answer which we shall study next time.

Suggested Reading

A very good general introduction to Differential Geometry can be found in the following books:

N J Hicks. Notes on Differential Geometry. Van Nostrand. 1965.

M Spivak. Differential Geometry. Vol. II. Publish or Perish, Berkeley. USA. 1970.

Address for correspondence
Kapil H Paranjape,
Indian Statistical Institute,
8th Mile, Mysore Road,
Bangalore 560 059, India.

Learning Organic Chemistry Through Natural Products

3. From Molecular and Electronic Structures to Reactivity

N R Krishnaswamy

N R Krishnaswamy was initiated into the world of natural products by T R Seshadri at University of Delhi and has carried on the glorious traditions of his mentor. He has taught at Bangalore University, Calicut University and Sri Sathya Sai Institute of Higher Learning. Generations of students would vouch for the fact that he has the uncanny ability to present the chemistry of natural products logically and with feeling.

In this part of the series, dynamic organic chemistry and organic reaction mechanisms are illustrated using the comparatively simple alkaloid papaverine.

As mentioned in the introduction to this series, (***Resonance***, Vol.1, No.1, 1996) the structure of a compound is like a milestone on a highway. What is more interesting, as pointed out by one of the giants in organic chemistry, Robert Robinson, is the surrounding countryside. In other words, it is not enough to know what a molecule looks like in a particular 'frozen' posture or profile. We need to find out what the molecule can do or be made to do. The most interesting chemical aspect of a molecule is its reactivity pattern.

The most interesting chemical aspect of a molecule is its reactivity pattern.

A structure is similar to a single photographic exposure of a subject. It is sufficient for identification but does not directly convey the complete character and the potential of the subject for action and reaction. It is true that an experienced and expert character reader may be able to get a great deal of significant

Robert Robinson was one of the all-time greats among organic chemists. This outstanding British chemist was a student of W H Perkin, Jr., who himself was trained by the great German chemist Adolf von Baeyer. Robinson once drew up a 'family tree' in which he traced his organic chemical training to Baeyer and from himself drew it down to some of his outstanding students which included two Indians: T R Seshadri and K Venkataraman. Among the other students of Robinson are Alexander Todd and Arthur Birch, names commonly encountered in organic chemistry text books.

information out of a single picture, but this is subjective. On the other hand, an album of photographs of a single subject taken at different times under different circumstances designed to bring out different moods can give a more comprehensive idea of the subject's character. In terms of molecular structure, what this means is a collection of structural formulae representing different possible conformations. It is also important to include details concerning the electronic structure of the molecule. One of the most efficient ways to accomplish this is by means of canonical (resonance) structures. The latter can be used to identify the electron rich and deficient regions of a molecule. Therefore, the information about electronic structure can be used to derive the nature of reactivity under different conditions.

It is not enough to know what a molecule looks like in a particular 'frozen' posture or profile. We need to find out what the molecule can do or be made to do.

Dynamic organic chemistry and organic reaction mechanisms can be profusely and effectively illustrated with examples from natural products. In this article, using the comparatively simple alkaloid papaverine as an example, the principles governing a few common organic reactions are highlighted, bringing out at the same time some of the subtle nuances which make them interesting and 'colourful'!

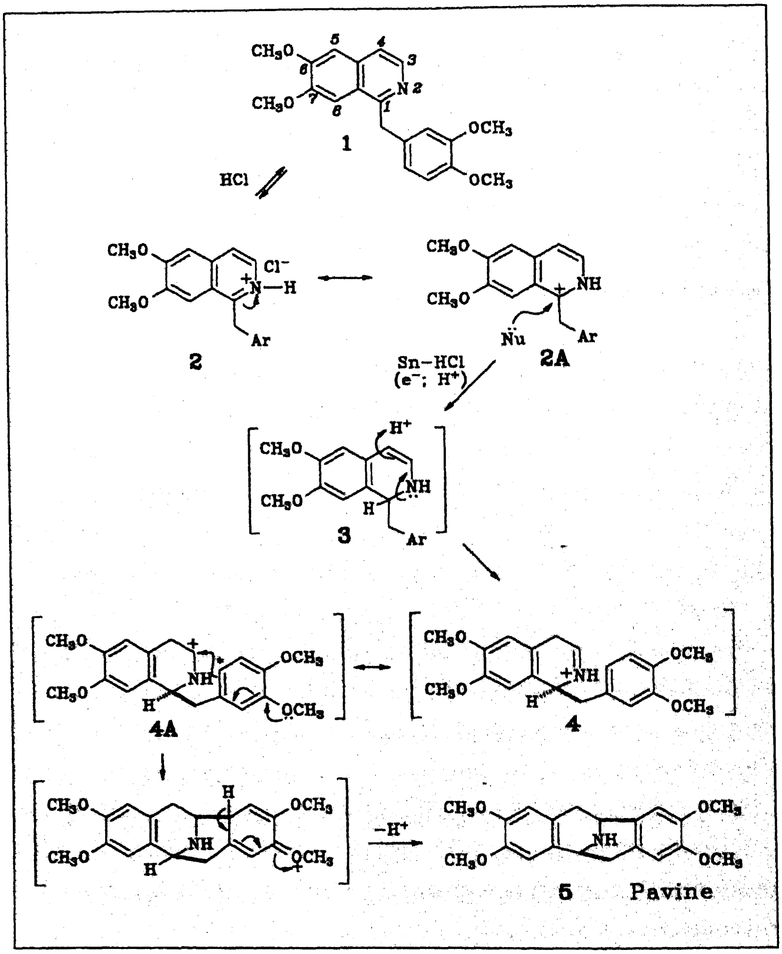
Papaverine (**1**) is one of the several nitrogenous basic constituents of opium[1] which is the dried latex of the unripe fruits of the plant *Papaver somniferum*. This compound is a (tetramethoxy)-1-benzylisoquinoline derivative. The compound is a mono tertiary base. Therefore, it readily forms a mono hydrochloride with HCl and reacts with a mole of methyl iodide to form a quaternary ammonium compound. For the mono hydrochloride (**2**), one can write an extreme canonical form (**2A**) which is a benzylic carbocation. Therefore, it can be attacked by any nucleophile, for example, an electron pair which can be supplied by a metal. Thus, the reduction of papaverine by tin and hydrochloric acid gives a dihydro derivative (**3**) in which the 1, 2-double bond of the pyridine part of the isoquinoline unit has undergone selective reduction. It is important to understand the principle involved in this selective reduction (as illustrated with structure **2A**) and to

[1] Opium contains a large number of alkaloids derived from the amino acid phenylalanine. The most important among these compounds is the narcotic analgesic morphine named after the Greek god of sleep, Morpheus. Robinson also called it a chemical Proteus, after another Greek god who could change his form, as the compound is susceptible to rearrangement.

compare this reaction with catalytic hydrogenation which, in this case, would not have been selective.

The resulting dihydro compound, being an enamine[2], then undergoes protonation, not on the nitrogen atom, but at position 4 as shown in *Scheme 1*. This is a consequence of the conjugation between the lone pair of electrons on the nitrogen atom and the double bond, and is similar to what one finds in a compound like pyrrole. The result is the formation of an iminium ion (**4**) for which one can write a carbocationic canonical form (**4A**). The latter then undergoes an intramolecular cyclization to yield a tetracyclic compound, known as pavine (**5**). As can be seen from the structural formulae, for this reaction to occur, the intermedi-

[2] The amino group of an enamine, like that of an amide, is either weakly basic or non-basic due to interaction between the orbital containing the lone pair and the π^* orbital of the neighbouring double bond. This makes the non-bonding electron pair less available for interaction with an extraneous acid. An extreme case is pyrrole in which this type of conjugation results in the generation of aromatic character which signifies molecular stability of a high order.



Scheme 1 Transformation of Papaverine to Pavine.

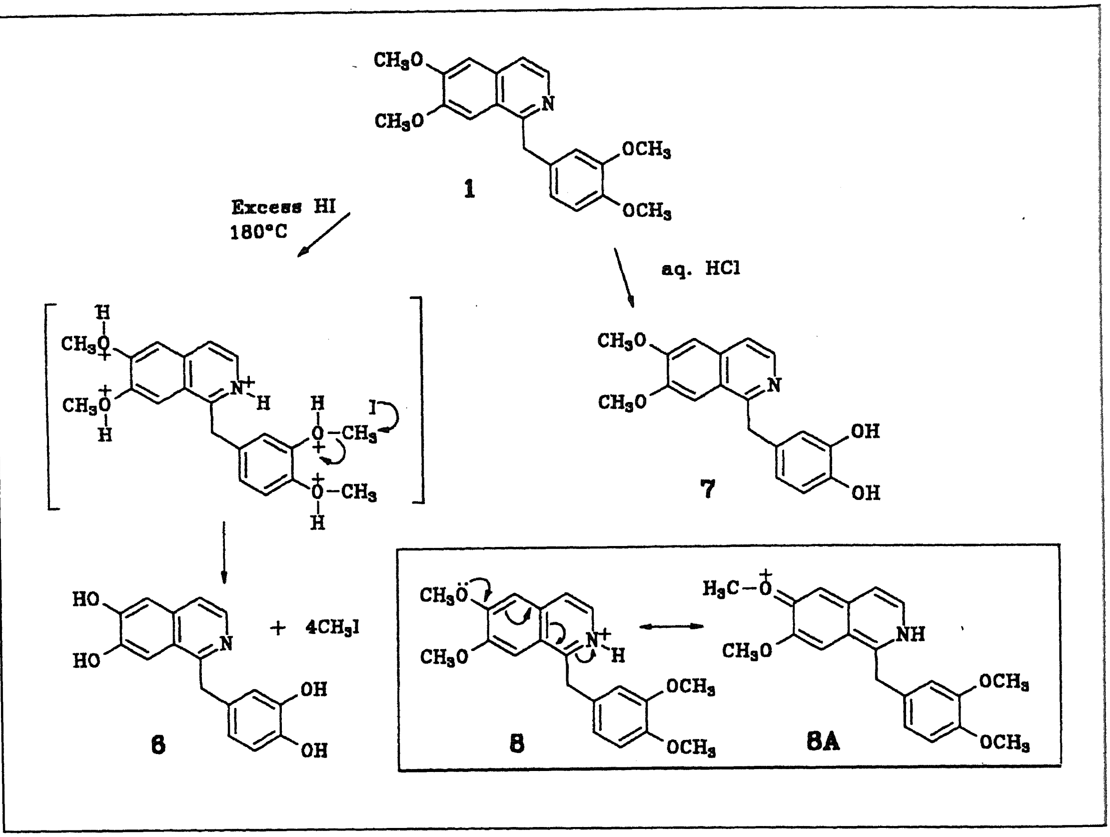
ate carbocation should be in the appropriate conformation, in which the side dimethoxyphenyl group can 'see' the reactive centre from a strategic angle and supply the electron pair needed for the formation of the new bond. The non-planar conformation of the partially reduced pyridine ring ensures proximity of the reactive centres which thus come within bonding distance.

The above example of a one pot reaction brings to light more than one fundamental principle which governs the course of organic reactions. The reaction can be classified as an intramolecular Friedel-Crafts alkylation[3] which is an aromatic electrophilic substitution reaction. The benzene ring which undergoes the 'alkylation' is activated by one of the two methoxyl groups which exerts a +M effect. The nitrogen containing ring loses its planarity when it undergoes reduction and this has the effect of raising the benzyl group at position 1 (i.e., the benzyl group can now occupy a pseudo-axial position). The carbon marked with an asterisk in structure **4A** can now come close to carbon 3 of the original isoquinoline unit. As in archery, so also in organic reactions, the trajectory of the attacking group is crucial. With the help of framework models, students should understand and appreciate this 'steric' aspect of organic reactions. Two-dimensional projectional formulae are inadequate for this purpose.

While the nitrogen atom is the chief basic centre in papaverine, the oxygen atoms of the four methoxyl groups, with their lone electron pairs, can also serve as Lewis bases. Thus, they can also be protonated with a strong acid. When heated with hydriodic acid, compound **1** undergoes demethylation to yield **6** as the final product (*see Scheme 2*). The methyl iodide generated in this reaction can be estimated after conversion into silver iodide. The Zeisel estimation of methoxyl groups is based on this principle and reaction.

Although the transformation in Scheme 2 is shown as a single reaction, it occurs in a step-wise manner. Interestingly, demethylation of the four methoxyl units does not follow a random

[3] The Friedel-Crafts alkylation reaction is one of the aromatic electrophilic reactions which find wide application in preparative organic chemistry (and industrial chemistry). The reaction is usually catalysed by a Lewis acid and the alkylating agent can be an alcohol, an aldehyde or an alkyl halide. In each case the reagent reacts with the Lewis acid to generate a carbocation which is the effective electrophile.



Scheme 2 Demethylation of Papaverine.

sequence. This is because the four groups are not equally basic. In other words, they do not undergo protonation with equal ease. The differences between them are revealed with the help of a weaker acid and under less vigorous conditions. The result is selective demethylation since demethylation is preceded by protonation. For instance, by heating with *aq*. HCl, papaverine can be demethylated to a dihydroxy compound (7) which still retains the two methoxyl groups on the isoquinoline unit. Under slightly more vigorous conditions, the methoxyl group at position 7 also undergoes demethylation. The oxygen atom of the methoxyl group which survives at position 6 is, therefore, the least basic of the four.

The above reactivity pattern can be understood in terms of the electronic structure of the molecule. The +M effect of the methoxyl group comes into play to different extents. The group at position 6 is in direct conjugation, through the intervening π

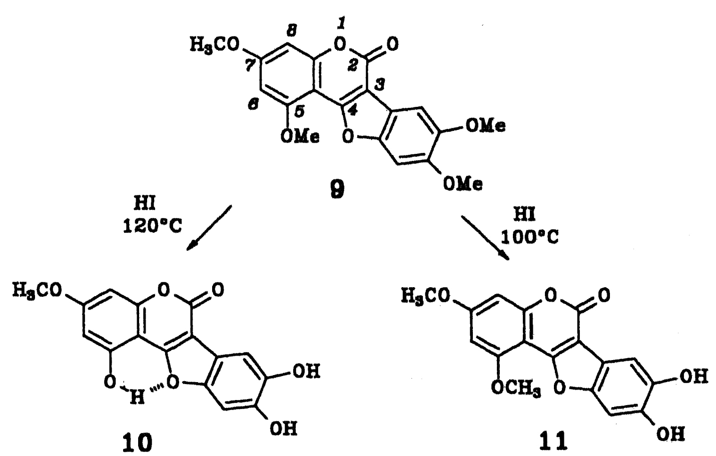
electrons, with the protonated nitrogen of the pyridine ring (see 8 and 8A in *Scheme 2*). In contrast, the methoxyl group at position 7 is not in conjugation with the protonated nitrogen, but the electron withdrawing effect of the isoquinolinium unit certainly lowers its basicity compared to the two methoxyl groups on the dimethoxybenzyl unit. These two methoxyl groups are 'insulated' from the protonated nitrogen, and are therefore free to react with an extraneous proton (i.e., they are more basic than the other two methoxyls).

Another example which reveals the subtle differences between differently located methoxyl groups is compound **9**, the trimethyl ether of wedelolactone[4] **10**. Like papaverine, this compound also has four methoxyl groups. But unlike papaverine, it is non-nitrogenous and is not an alkaloid. However, its behavior towards acidic demethylating agents is similar. In this case also selective demethylation under controlled conditions gives a product which retains two of the methoxyl groups to produce **11**. Under more vigorous conditions a third one is lost and the product is wedelolactone as shown in *Scheme 3*.

In this case, the different methoxyl groups are differentiated by the presence of the carbonyl group of the lactone, which, like the

[4] Wedelolactone was first isolated by T R Govindachari and his co-workers from the Indian medicinal plant, *Wedelia calendualacea* which is known as Bhringa raj in Sanskrit. The leaves of this plant are traditionally used for the treatment of jaundice and other liver disorders. Like T R Seshadri and K Venkataraman, T R Govindachari has also made significant contributions to the chemistry of natural products. He had his initial training from the American chemist Roger Adams.

Intramolecular hydrogen-bonding can significantly influence ground-state properties as well as reactivities of organic compounds.



9

10

11

Scheme 3 Demethylation of wedelolactone.

What matters is not the label but the fundamental forces responsible for the building and demolition of molecules!

protonated nitrogen of papaverine is electron-attracting. The methoxyls at positions 5 and 7 are in direct conjugation with this carbonyl group and, therefore, are not sufficiently basic to undergo protonation at the rate necessary for demethylation under mild conditions. On the other hand, the other two methoxyls are not in direct conjugation with the lactone carbonyl. Hence, protonation and demethylation occur readily at these positions. Between the methoxyl groups at positions 5 and 7, protonation at the former is favored. This is because of the possibility of a strong hydrogen bond with the neighbouring oxide bridge. The same stabilising interaction is present in the phenol formed in the reaction (**10**). Intramolecular hydrogen-bonding can significantly influence ground-state properties as well as reactivities of organic compounds.

Taken together as a complementary pair, papaverine and tri-O-methylwedelolactone bring out the subtle relationships between the structural parameters and the mechanism of acid-catalysed dealkylations of phenolic ethers. When studied this way one can clearly perceive the few controlling principles which govern the behavior of organic molecules, whether synthetic or natural, and irrespective of whether a compound is an alkaloid or a flavonoid or a terpenoid. What matters is not the label but the fundamental forces responsible for the building and demolition of molecules! It is like looking into a kaleidoscope which, with each shake, gives a different pattern with the same number of glass pieces.

Address for correspondence
N R Krishnaswamy,
Visiting Professor,
NMKRV College for Women
with PG Centre,
Jayanagar III Block,
Bangalore 560011, India.

Albert Einstein ... who refused to believe that "God plays dice with the world" writing to Niels Bohr in 1924 said: "I cannot bear the thought that an electron exposed to a ray should by its own free decision choose the moment and the direction in which it wants to jump away. If so, I'd rather be a cobbler or even an employee in a gambling house than a physicist".

River Piracy

Saraswati that Disappeared

K S Valdiya

The legendary river Saraswati, which flowed from the Himalaya and emptied finally into the Gulf of Kachchh, has vanished. Tectonic movements change river courses, behead streams and sometimes even make large rivers such as the Saraswati disappear.

K S Valdiya is at Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore.

Mighty River of Vedic Time

There was this highly venerated river Saraswati flowing through Haryana, Marwar and Bahawalpur in Uttarapath and emptying itself in the Gulf of Kachchh, which has been described in glowing terms by the *Rigveda*. "Breaking through the mountain barrier", this "swift-flowing tempestuous river surpasses in majesty and might all other rivers" of the land of the pre-Mahabharat Vedic

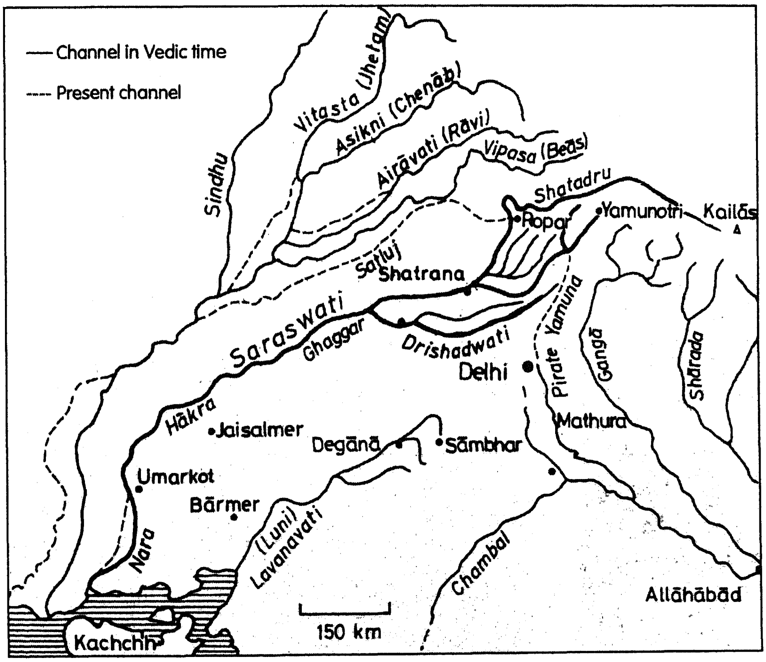
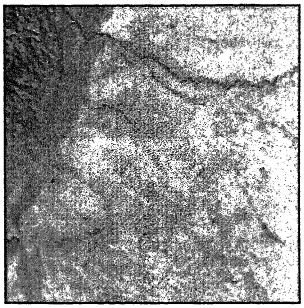


***Figure 1** (bottom left) Legendary Saraswati of the Vedic times was formed by joining together of the Shatadru (Satluj) and what is today known as the Yamuna. The Aravali was not a highland but a thickly forested terrain sloping southwestwards.*

period. More than 1200 settlements, including many prosperous towns of the Harappan culture (4600 to 4100 years Before Present - BP) and ashrams of *rishis* (sages) lay on the banks of this life-line of the Vedic time.

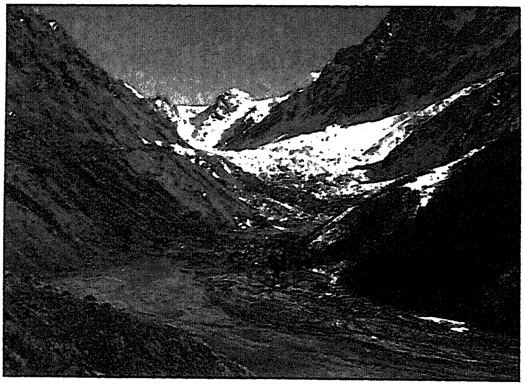
Figure 2 (bottom) Satellite picture of the Haryana-Punjab region, showing the disproportionately wide channels (with little or no water) abandoned by big rivers which have migrated to the east or west.

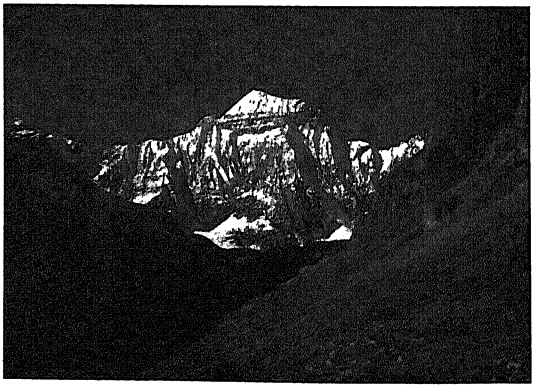


Where has that great river gone? It is today represented by the disproportionately wide and astonishingly water-less, sand-filled channels of Ghaggar in Haryana and Marwar, Hakra in adjoining Bahawalpur, and Nara in Sindh (*Figure 1*). These channels, which discharge only floodwaters, are quite apparent in the satellite imageries (*Figure 2*).

The legendary Saraswati was indeed a great river which rose in the Bandarpunch massif of the Great Himalaya in western Garhwal (*Figure 3A*), flowed southwestward through a channel past Adibadri, Bhavanipur and Balchhapur in the foothills, and met the Shatadru or Satluj (which then veered towards the southeast). The Shatadru came from the region of Mount Kailas in southwestern Tibet (*Figure 3B*). The ancient Saraswati was thus formed by the confluence of what are today the Yamuna and Satluj rivers flowing in entirely different directions (*Figure 1*). The two joined at Shatrana, 25 km south of Patiala, and flowed through a 6 to 8 km wide channel (*Figure 4*) known today as the Ghaggar. Obviously, a large volume of water flowed down the Ghaggar channel. Even today the combined discharge of the Yamuna and Satluj is of the order of 2900 million cubic metres per year. It must have been many times more in those days.

Figure 3(A) (left): Main confluent of the Saraswati, what is today called the Tons branch of the Yamuna, springs from the Har-ki-Dun glacier. (B) (right): Source of the other confluent of the Saraswati, the Satluj, lies beyond the Indo-Tibetan border range.





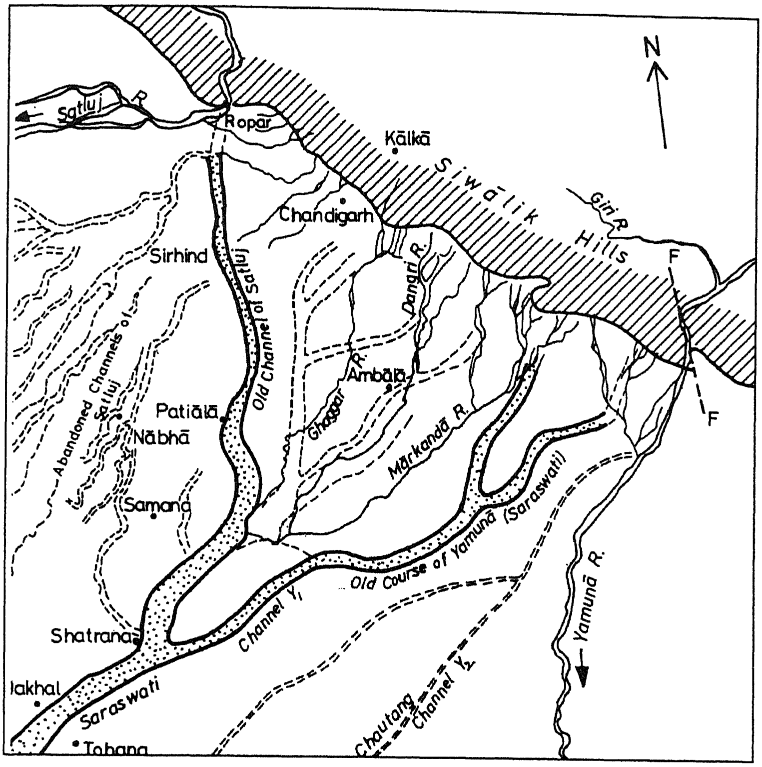


Figure 4 Dry channels of the Ghaggar and its tributaries seem to have been the former courses of the confluents of the Saraswati. (Based on Yashpal et al, 1980).

The Ghaggar is known as Hakra in northwestern Marwar and Bahawalpur (Pakistan) and as Nara in Sindh, before it discharges into the Gulf of Kachchh. Drishadwati - now a dry channel called Chautang - joined the Saraswati near Sirsa from the east (*Figure 1*). It was at Kurukshetra in Manu's Brahmavarta between the Saraswati and the Drishadwati where the epic battles of Mahabharat were fought in the post-Vedic period.

It was at Kurukshetra in Manu's Brahmavarta between the Saraswati and the Drishadwati where the epic battles of Mahabharat were fought

Wetter Period in Marwar

Western Rajasthan - including the Thar tract - was a wetter region some 40,000 years ago. Periods of dryness alternated with phases of wetness. This is testified by pollen grains buried and trapped in the sediments of the Lunkaransar and Didwana lakes and by thermoluminescence of sands in dunes and floodplains. The Saraswati and its tributaries held sway in the northern part, and the Lavanavati (Luni) had an organized drainage network of perennial streams in the southern part. It was in this well-watered,

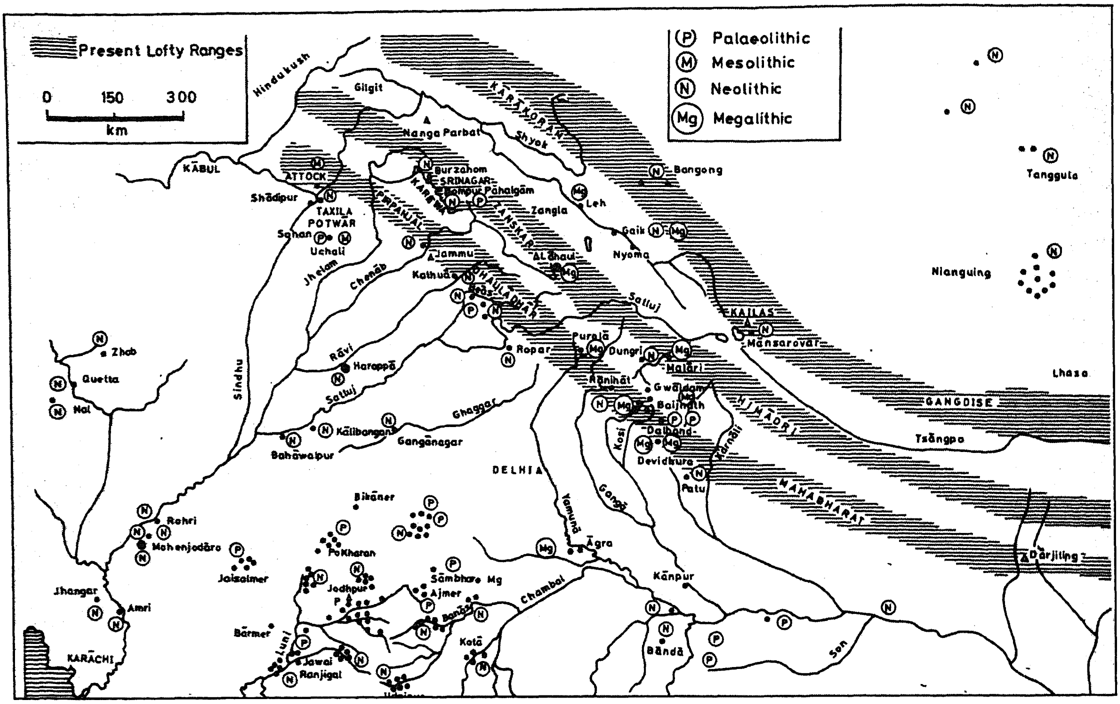


Figure 5 Site of settlements of the stone-age people in the Palaeolithic to Neolithic period (Based on V N Misra 1995 and other sources).

presumably fertile and congenial land of the Saraswati, Drishadwati and Luni that the stone – age people established their settlements (*Figure* 5), and developed their Palaeolithic, Mesolithic and Microlithic cultures.

Stone –age settlements evolved and developed their Palaeolithic, Mesolithic and Microlithic cultures in the well-watered, presumably fertile and congenial land of the Saraswati, Drishadwati and Luni.

From 10,000 to 3,500 years BP, the climate was quite wet - the rainfall being almost three times what it is now (*Figure* 6). This is indicated by the analysis of pollen (dominated by those of *Syzygium, Pinus* and *Astemisia*). Cutigens in pollens and fragments of charcoal of stubbles imply that these people had taken to agriculture - 9,400 years BP in the area of the Lunkaransar and 8,000 years BP in the Sambhar lake tract.

More than 75% of the 1,600 settlements of the Harappan culture have been found in the valley of the Saraswati, such as at Banawali and Kalibangan in the Ghaggar valley and AliMurad and Kot in the Hakra valley. The Harappan civilization, dating back to the period 4,600-4500 to 4,200-4,100 years BP, was spread over nearly 13 lakh square kilometre area, stretching from Sutkongedar in

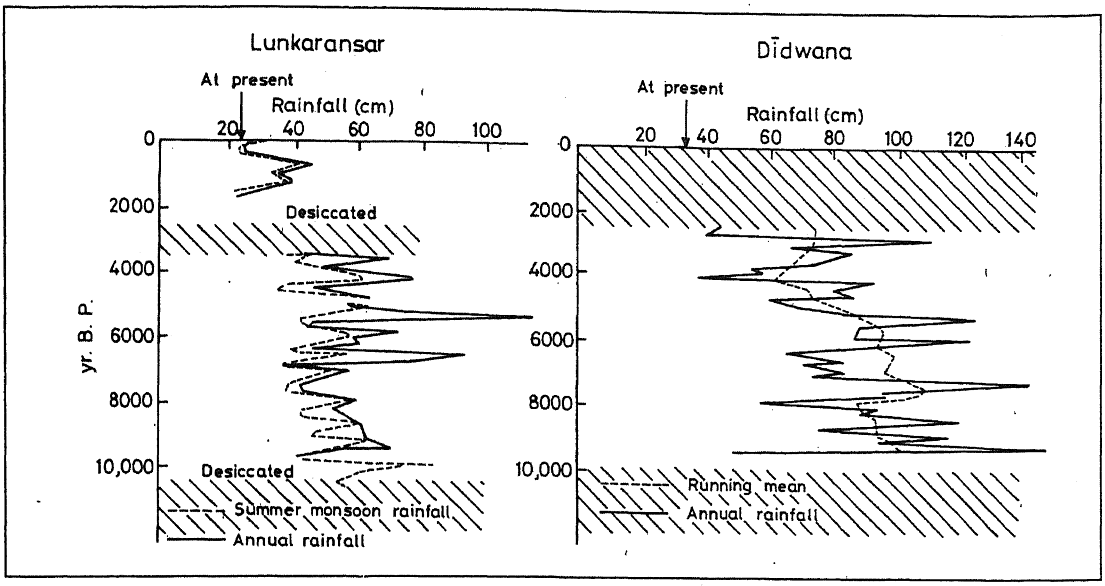
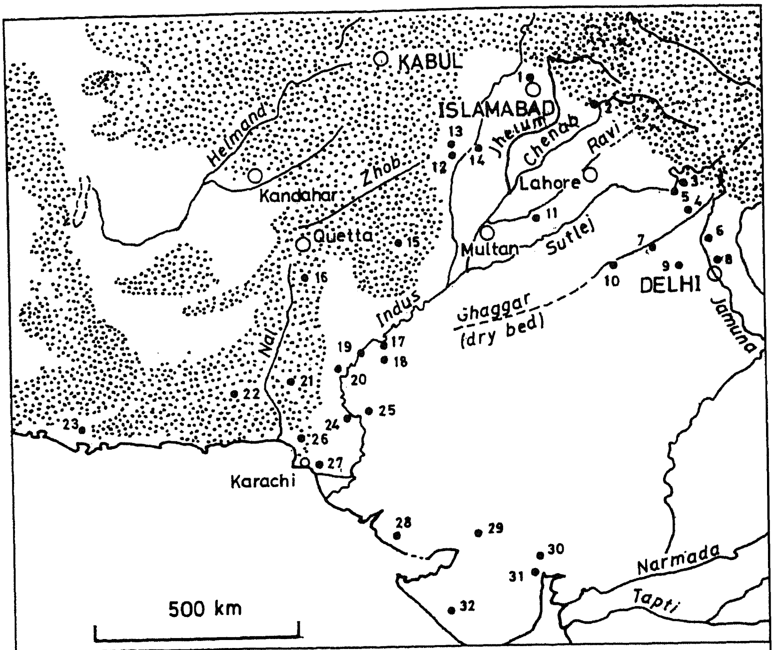


***Figure 6** Analyses of pollen buried with sediments in the Lunkaransar and Didwana lakes indicate — according to Gurdip Singh and his colleagues (1974) — that in the period 10,000 to 3,500 years BP Western Rajasthan used to have at least three times the rainfall that it has today.*

the west, through Mohenjodaro in the westnorthwest, Ropar in the north, Alamgir in the east, Sutkotri in the south to Lothal, Rangpur, Rojri and Dhaulavira in the southwest (*Figures* 5 and 7). The older Harappan sites are concentrated in the lower reaches of the Saraswati, while later Harappan settlements nestle in its upper reaches - in the Siwalik domain. There seems to have been upstream migration around 3,700 years BP, presumably prompted by a decline in the river discharge. Why was there a reduction in the river discharge? Perhaps the climate had worsened, as indicated by the lake waters turning saline around 3,700 years BP (borne out by overwhelming appearance of halphytes among the aquatic flora of the lakes). Or, perhaps the Saraswati had been robbed of its water.

Ganga Stole Away Saraswati's Water

Tectonic movements overtook the northern part of the Indian subcontinent, and the Aravali started slowly rising . The evidence for the continuing rise or uplift of the Aravali Range is quite striking. The western flank delimited by faults is marked by very steep straight scarps. The gently west-flowing streams draining the very old mature terrain of Mewar either descend suddenly in



1	Sarai Kala	2	Manda	3	Rupar	4	Bhagwanpura
5	Bara	6	Hulas	7	Banawall	8	Alamgirpur
9	Mitathal	10	Kalibangan	11	Harappa	12	Gumla
13	Rahman Dheri	14	Musa Khel	15	Dabar Kot	16	Nausharo
17	Sukkur	18	Kot. Diji	19	Mohenjo-daro	20	Lohamjo-daro
21	Nindowari	22	Kulli	23	Sutkagen Dor	24	Amri
25	Chanhu-daro	26	Balakot	27	Allahdino	28	Desalpur
29	Surkotada	30	Lothal	31	Rangpur	32	Rojdi

***Figure 7** Major settlements of the Harappan period (A H Dani and B K Thapar).*

waterfalls, or flow through deep gorges and ravines in the western flank of the range. These streams are characterized by entrenched meanders and incised channels, and show development of uplifted terraces on their banks before abruptly swerving across the active faults.

The Ganga had robbed the Saraswati of the major portion of its water through the agency of a branch of its tributary, the Chambal.

The Saraswati was forced to shift its course — progressively eastward. The Chautang channel (*Figure 1*) possibly represents the course abandoned by the eastward migrating Saraswati. Uplift of the Aravali domain accentuated the pace of erosion of the terrain. Consequently a branch of the Chambal River started

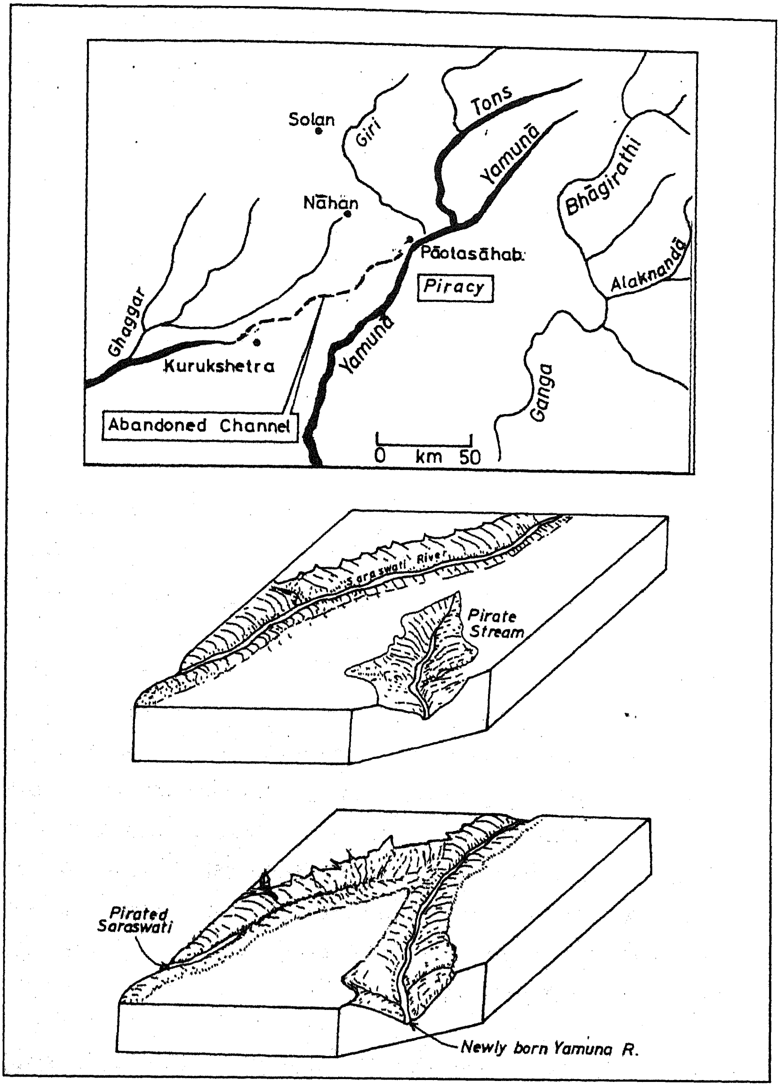


***Figure 8** A south-flowing branch of the Chambal, (the southwestern tributary of the Ganga) cut its channel headwards and captured the water of the then south-west flowing Saraswati. The new channel, through which the diverted water flowed, was later named Yamuna. Map shows the drainage pattern after this river piracy — after the Saraswati was robbed of its water by the Ganga through the agency of its tributary.*

cutting its course northwards by headward erosion. It cut the channel deeper than that of the Saraswati (*Figure 8*), and thus beheaded the Saraswati. During rains, the floodwater of the Saraswati rushed into this new channel (later to be called Yamuna) culminating in the capture of the Saraswati by the Chambal, the southwestern tributary of the Ganga. This was a case of river piracy, resulting from accelerated headward erosion, which in turn was prompted by tectonic uplift of the terrain. Thus the Ganga had robbed the Saraswati of the major portion of its water through the agency of a branch of its tributary, the Chambal (*Figure 8*).

Deprived of the snow–fed waters of the Yamuna and the Satluj, the Saraswati was reduced to a puny river, left with the streams rising in the Siwalik domain.

The reduced flow in the Saraswati, coupled with the onset of dry climatic conditions over western Rajasthan, forced the Harappans to migrate upstream and settle down in the foothills of the Siwalik domains. This must have happened about 3,700 years BP The *Markandeya* and the *Varaha Puranas* tell us that the Saraswati was in decline during the Mahabharat time. Sage Manu states that the Saraswati vanished in the sand at Vinasan, near Sirsa. There is allusion to the disappearance of the river in *Van Parva* of the *Mahabharat,* and also in the *Siddhant Shiromani.*

Great Betrayal

The Satluj joined the Sindhu, and the Saraswati was left high and dry, having being betrayed once again.

The Aravali continued to rise. The newly formed Yamuna was forced to migrate progressively eastward. Satellite imageries show that it has migrated 10 to 40 km (in different segments) since the time of Lord Krishna, who was born in a prison on the bank of the Yamuna. The Satluj likewise moved westward, abandoning its older channels successively. Dry channels such as Wah, Naiwal and Sarhind bear testimony to the progressive westward shifting of the Satluj. Finally it got deflected, possibly as a result of paroxysmal uplift of the Aravali domain and concomitant subsidence of the land to the west. This is obvious from the spectacular U-turn of the Satluj at Ropar (*Figure 9*). The Satluj joined the Sindhu, and the Saraswati was left high and dry. Saraswati was betrayed once again.

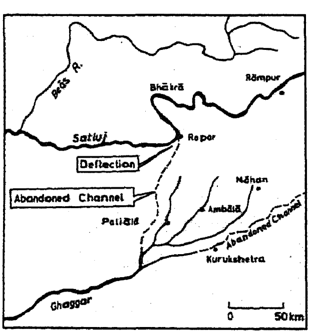


Figure 9 Great betrayal. Later when the Aravali rose, and as the land to the west sank, the Satluj changed its course abruptly, making a sharp U-turn at Ropar.

However, some water of this Himalayan river continued to flow into the Hakra-Nara channel until about 1245 AD, when there was a great migration of the desert people out of the region. The Satluj finally ceased to contribute water in 1593 AD, when it changed its course finally and decisively.

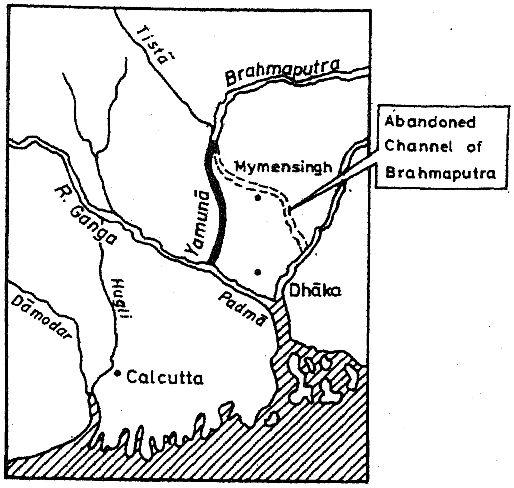
Deprived of the waters of the two snow-fed rivers (Yamuna and Satluj), the Saraswati was reduced to a puny river, left with the waters of the petty streams rising in the Siwalik domain - Wah, Ghaggar, Dangri, Markanda, Sarsuti etc. Only flood waters flowed down the large channel that was once the mighty Saraswati.

Rivers Named Yamuna

The Saraswati of the Vedic period was beheaded and robbed of its water by a branch of the Chambal, a tributary of the Ganga. The channel through which the stolen water flowed is known as Yamuna.

Skirting the Meghalaya massif the Brahmaputra used to flow east through Mymensingh in Bangladesh to meet Meghana. Then the Barind terrain started rising between 1720 and 1830 A.D., and the Brahmaputra was attracted towards the Ganga. Abandoning its old course and the Meghana, the Brahmaputra joined the Ganga, west of Dhaka. The new pirate river is called Yamuna.

Between the Meghalaya and Mikir hills in Assam flows Kapili, merging with Brahmaputra south-west of Naugaon. A branch of the neighbouring Dhanshree captured its headwaters. The new channel is named Yamuna.



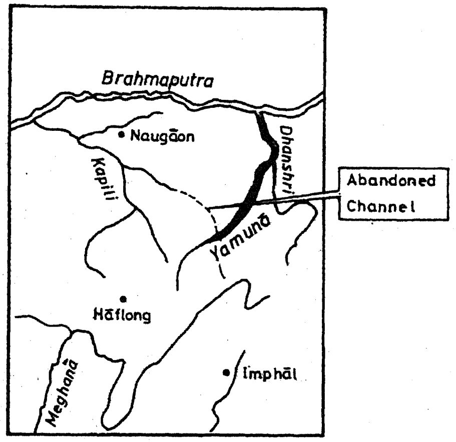


Figure 10 In Assam and Bangladesh, as in Haryana - UP, the channel through which river piracy occurred, is named Yamuna.

Western Rajasthan gradually turned into a parched land of moving sands.

It was not only the Satluj that was moving westwards. Indeed all the rivers of the Sindhu system—including the Asikni (Chenab), the Vipasa (Beas) and the Sindhu itself have been shifting perceptibly. The Sindhu migrated 160 km westwards in historical times. It appears that the uplift or rise and subsidence or sinking of the ground resulting from crustal movements causes changes

in the courses of rivers, the beheading of streams, the piracy of their waters, and the disappearance of rivers, some even as great as the River Saraswati. This is the effect of the continuing tectonic subsidence of the belt adjoining the Pakistani mountain front.

The Saraswati is no more. But the anastomosing network of dry channels which lose themselves in the desert sands, tells us of the river that was great, and of the human history which was glorious. The network of canals across several states implies the return of the Saraswati to the land that was once very green and fertile.

The Saraswati is no more. But the anastomosing network of dry channels which lose themselves in the desert sands, tells us of the river that was great, and of the human history which was glorious.

Suggested Reading

R D Oldham. On probable changes in the geography of the Panjab and its rivers, *J. Asiatic Soc. of Bengal*, 55:322-343. 1886.

C F Oldham. The Saraswati and the lost river of the Indian desert, *J. R. Asiatic Soc.*, 34:49-76. 1893.

S C Sharma. The description of rivers in the Rigveda, *The Geographical Observer*, 10:79-85. 1974.

G Singh, R D Joshi, S K Chopra A B Singh. Late Quaternary history of vegetation and climate of Rajasthan desert, India, *Philos. Trans. R. Soc. London*, 267B:467-501. 1974.

B Ghose , Amal Kar, A Husain. The lost courses of the Saraswati river in the great Indian deserts; new evidence from landsat imagery, *Geographical J.*, 145:446-451. 1979.

Yashpal, B Sahai, R K Sood, D P Agrawal. Remote sensing of the lost Saraswati river, *Proc. Indian Acad. Sci. Earth Planet. Sci.*, 89: 317-331. 1980.

Amal Kar, B Ghose. The Drishadvati river system of India: an assessment and new findings. *Geomorph. J.* . 150:221-229. 1984.

A K Singhvi, A Kar. Thar Desert in Rajasthan. Geological Society of India, Bangalore. 191. 1992.

V N Misra. Geoarchaeology of Thar Desert, Northwest India, in: S Wadia et al, (Eds) Quaternary Environments and Geoarchaeology of India. Geological Society of India, Bangalore. 210-230. 1995.

Address for correspondence
K S Valdiya,
Jawaharlal Nehru
Centre for Advanced
Scientific Research,
Bangalore 560 064, India.

Hardy on Ramanujam Theorems ... "They defeated me completely, I had never seen anything in the least like them before. A single look at them is enough to show that they could only be written down by a mathematician of the highest class. They must be true because, if they were not no one would have had the imagination to invent them".

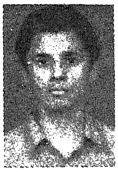
Chaos Modelling with Computers

Unpredictable Behaviour of Deterministic Systems

Balakrishnan Ramasamy and T S K V Iyer

Chaos is a type of complicated behaviour found in non-linear dynamical systems. Computers are playing an important role in the growth of this science.

Chaos is one of the major scientific discoveries of our times. In fact many scientists rank it along with relativity and quantum mechanics as one of the three major scientific revolutions of this century. Chaos is a science of everyday things: it has been implicated in areas ranging from heart failure, meteorology, economic modelling, population biology to chemical reactions, neural networks, fluid turbulence and more speculatively even manic-depressive behaviour. It also seems to occur everywhere — in rising columns of cigarette smoke, in fluttering flags, in dripping faucets, in traffic jams and so on. Computers have played a major role in the discovery and subsequent developments in this field. The computer is to chaos what cloud chambers and particle accelerators are to particle-physics. Numbers and functions are chaos' mesons and quarks. In this article we provide an introduction to chaos and the role that computers play in this field.



Balakrishnan Ramasamy is an electrical and electronics engineer who works in Siemens Communication Software Ltd., Bangalore as a software engineer.

T S K V Iyer is a visiting professor at the Birla Institute of Technology and Science, Pilani. He has been at BITS in various capacities and was professor of EEE and deputy director when he retired in 1987. He is involved with educational innovations, special student projects and conducts a workshop for new faculty.

Chaos and Dynamical Systems

The laws of science aim at relating cause and effect and thereby making predictions possible. For example, based on the laws of gravitation, eclipses can be predicted thousands of years in advance. But there are other natural phenomena that are not predictable though they obey the same laws of physics. The weather, the flow of a mountain stream, the roll of a dice are examples of such phenomena. It was believed until recently that precise predictability can in principle be achieved, by gathering and processing sufficient amount of information.

Simple deterministic systems can generate random like behaviour. The randomness is fundamental; gathering more information will not make it go away. Randomness generated in this way has come to be called *chaos*.

Such a viewpoint has been altered by a striking discovery: simple deterministic systems can generate random like behaviour. The randomness is fundamental; gathering more information will not make it go away. Randomness generated in this way has come to be called *chaos*.

The discovery of chaos has created a new paradigm in scientific thinking. On the one hand it places fundamental limits on the ability to make predictions. On the other hand, the determinism inherent in chaos implies that many random phenomena are more predictable than had been thought. Chaos allows us to find order amidst disorder. The result is a revolution affecting many different branches of science.

Chaos has put an end to the Laplacian fantasy of mechanistic determinism. Laplace once boasted that "The present state of the system of nature is evidently a consequence of what it was in the preceding moment, and if we conceive of an intelligence which at a given instant comprehends all the relations of the entities of this universe, it could state the respective positions, motions, and general affects of all these entities at any time in the past or future." But according to chaos, even if we were able to write down the equations governing every particle in the universe, if we knew the state of the system at any point of time *only* approximately, we cannot predict the behaviour of the system in the long term because the small errors inherent in measurements amplify very fast, thus making prediction impossible. This property of *sensitive dependence on initial conditions* is one of the characteristics of chaos.

The property of *sensitive dependence on initial conditions* is one of the characteristics of chaos.

Chaos emerges from the larger context of the theory of dynamical systems. A *dynamical system* is one that allows us to predict the future given the past. A dynamical system is made up of two parts: the notion of a phase (the essential information about a system) and a dynamic (a rule that describes how the state evolves with time). The evolution of the system can be visualized in a phase

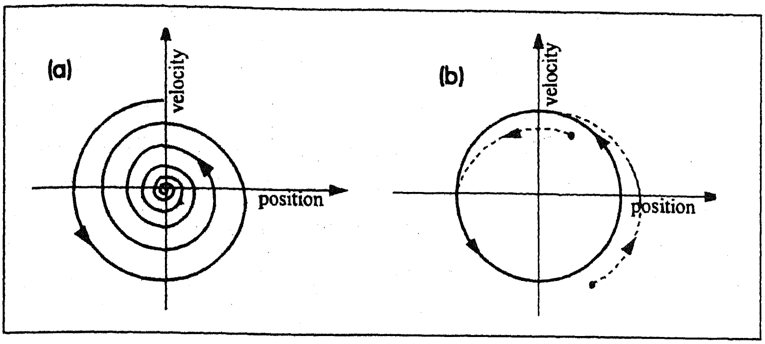


Figure 1 (left) Phase space portrait for a simple damped pendulum has a fixed point attractor; (right) phase space portrait for a pendulum clock shows a limit cycle. Any arbitrary initial condition (represented by the dots) is attracted to the limit cycle.

space, an abstract construct whose coordinates are the components of the state.

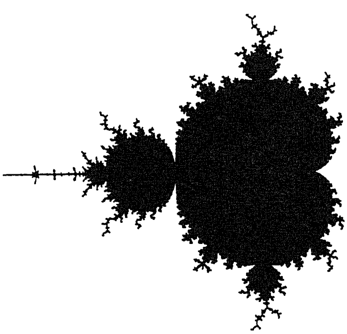
The simple pendulum is a good example of a dynamical system. The position and velocity are all that are needed to determine the motion of a pendulum. The phase is thus a point in a plane, whose coordinates are position and velocity. Newton's laws provide the dynamic (or rule), that describes how the phase evolves. As the pendulum swings the phase moves along an orbit in a plane. For an undamped pendulum the orbit is a loop whereas for a damped pendulum the orbit spirals to a point called the *fixed point* as the pendulum comes to rest.

The example of a pendulum also introduces us to the concept of an *attractor*. An attractor is a finite region in phase space, to which the system settles down in the long run. For example the fixed point is an attractor for a simple damped pendulum (*Figure 1a*). It can be thought of as attracting all the nearby points to itself. Another type of attractor could be found in the behaviour of a pendulum clock. Since the energy lost due to friction is balanced by the energy input to the system, the pendulum executes periodic motion. The phase space portrait of a pendulum clock is a cycle. Irrespective of how the pendulum is set swinging, it approaches the same cycle in the long-term limit. Such attractors are called *limit cycles* (*Figure 1b*). Systems whose attractors are classical[1] such as fixed points and limit cycles, have the property that small measurement errors remain bounded and long term behaviour is predictable.

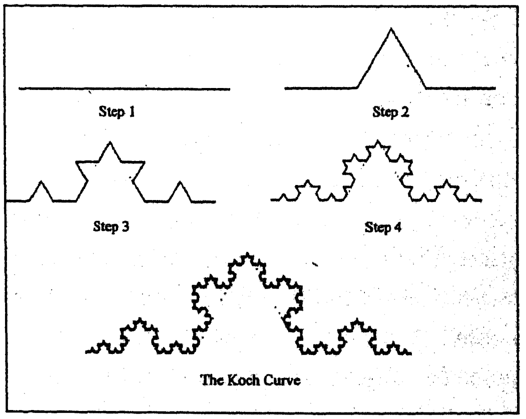
[1] Most systems settle down in the long run to a fixed point or a limit cycle. These classical attractors remain bounded in spite of small variations in initial conditions.

Fractals : A Geometry of Nature

"Clouds are not spheres; mountains are not cones; coastlines are not circles and bark is not smooth, nor does lightning travel in a straight line", Benoit B Mandelbrot, the discoverer of fractals is fond of saying. Traditional Euclidean geometry is insufficient in describing natural objects such as clouds, mountains, coastlines, etc. To explain such natural shapes Mandelbrot invented fractal geometry.Fractals have the property of self-similarity, i.e. parts of the fractal object resemble the whole across all scales of magnification. In other words fractals display symmetry across scales. Fractals provide the mathematics necessary to describe the phase space portrait of chaotic systems. In general, fractal objects have a fractional dimension unlike traditional Euclidean shapes whose dimension is an integer. However, there are exceptions. The well-known Mandelbrot set has a fractal boundary with a fractal dimension equal to 2, the dimensions of an Euclidean plane. Interestingly this important fact was established only recently, in 1991, by Mitsuhiro Shishikura. Another fractal object which has a fractal dimension of 2 is the so called 'Skewed Web'. This is a three dimensional object which is an analogue of the famous Sierpinski's Gasket. In this context it may also be mentioned that the network of blood vessels is not only a fractal but is said to have a fractal dimension equal to 2. (Refer articles by Ian Stewart given in Suggested Reading).



The Koch curve is a good example of a fractal. The curve is constructed as follows. A line segment is taken and is divided into three parts. The middle one-third of the line segment is replaced by two line segments as shown in Step 2 of the figure. The above operation is now performed on the four line segments (Step 3). This process is carried on ad-infinitum. At the end of the process we get a fractal object called the Koch curve. The Koch curve has a dimension of 1.261... This is because the Koch curve is neither a line (dimension 1) nor a plane (dimension 2), but something in-between. It can be observed that the Koch curve is self-similar.



Generating the Koch curve starting from a line segment

The Discovery Of Chaos

Edward Lorenz, a meteorologist at MIT first discovered chaos in the early 60's. Lorenz wrote a system of equations that modelled the earth's weather and simulated its behaviour on his Royal MacBee computer. His computer spewed out a series of numbers, which represented various weather parameters of his model. One day wanting to examine a sequence at greater length, he took a short cut. Instead of starting the whole run again, he started mid-way through, by typing in the numbers from an earlier printout. When he examined the results, he discovered that his new weather patterns were diverging very rapidly from the patterns of the last run. Lorenz found out that the problem lay in the numbers he had typed. In the computer's memory, six decimal places were stored whereas only three appeared on the printout. An error of one part in a thousand had changed his weather patterns drastically. Lorenz called his discovery "the butterfly effect" - the notion that a butterfly flapping its wings in Bombay will set off a tornado in Japan a week later. Technically, the butterfly effect is called sensitive dependence on initial conditions, which is one of the hallmarks of chaos.

The advent of chaos introduces us to a new type of attractor - a *strange attractor* or a chaotic attractor. Geometrically a strange attractor is a *fractal* (see box on fractals), i.e. it reveals more detail as it is increasingly magnified. In strange attractors, arbitrarily nearby orbits diverge exponentially fast and so stay together for only a short time. Strange attractors (and hence chaos) are found in certain non-linear dynamical systems. Non-linear systems, unlike their linear counterparts, do not have closed form solutions. Hence one has to numerically simulate the behaviour of the non-linear system. Therefore, it is not surprising that the growth in the field of chaos has occurred hand-in-hand with the widespread availability of computing power.

The original discovery of chaos, by meteorologist, E N Lorenz, (see box on discovery of chaos) in the following set of differential equations, is a good example of the role that computers play in chaos.

$$\begin{aligned} dx/dt &= 10y - 10x \\ dy/dt &= -xz + 28x - y \\ dz/dt &= xy - (8/3)z \end{aligned} \tag{1}$$

The advent of chaos introduces us to a new type of attractor - a *strange attractor* or a chaotic attractor. Geometrically a strange attractor is a *fractal* .

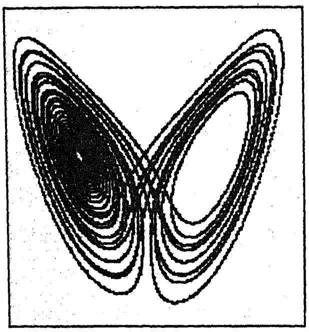


Figure 2 Trajectory of the Lorenz attractor in the x-z axis, obtained by plotting equations (1).

The trajectory of this system of equations projected on the x-z plane, is called the *Lorenz attractor,* (*Figure* 2). The books by Gleick and Gulick discuss this attractor in greater detail.

Simplicity and Universality in Transition to Chaos

Chaos occurs even in the most deceptively simple systems. For example, it was observed that the following equation, called the *logistic equation* (see Robert May's article in *Nature*) captured the essence of Lorenz's system of equations.

$$x_{n+1} = \mu . x_n (1 - x_n) \qquad (2)$$
where $0 < x_n < 1$ and $1 < \mu < 4$.

The above equation is also a dynamical system, but the evolution of the system is in discrete time instead of continuous time. An initial value for x_n is chosen and is called x_0. Equation (2) then gives us the value of x_1. This simple calculation is repeated endlessly, feeding the output of one calculation as input for the next. (In computer parlance this process is called iteration.) In analysing the equation we are interested in finding out the behaviour of x_n as $n \to \infty$. In actual practice, we observe the behaviour of x_n after a few thousand iterations (to allow the transients to die down).

Non-linear systems, unlike their linear counterparts, do not have closed form solutions. Hence one has to numerically simulate the behaviour of the non-linear system. Therefore, it is not surprising that the growth in the field of chaos has occurred hand-in-hand with the widespread availability of computing power.

When the parameter μ is less than three, for any initial condition (i.e. the value of x_0) between 1 and 3, the value of x_n converges to a fixed point or steady state (*Figure* 3a). As the value of the parameter μ is increased beyond three, the limit value of x_n oscillates between two values. For example, for a μ value of 3.2, the value of x_n keeps oscillating between the values 0.7994 and 0.5130. The system is said to have a period of two. The time-evolution of a period two system is depicted in the diagram in *Figure* 3b. At the parameter value of 3, the system is said to have undergone a *period-doubling bifurcation,* because its behaviour changes from steady state to that of period two.

As the parameter value is increased further, the system undergoes

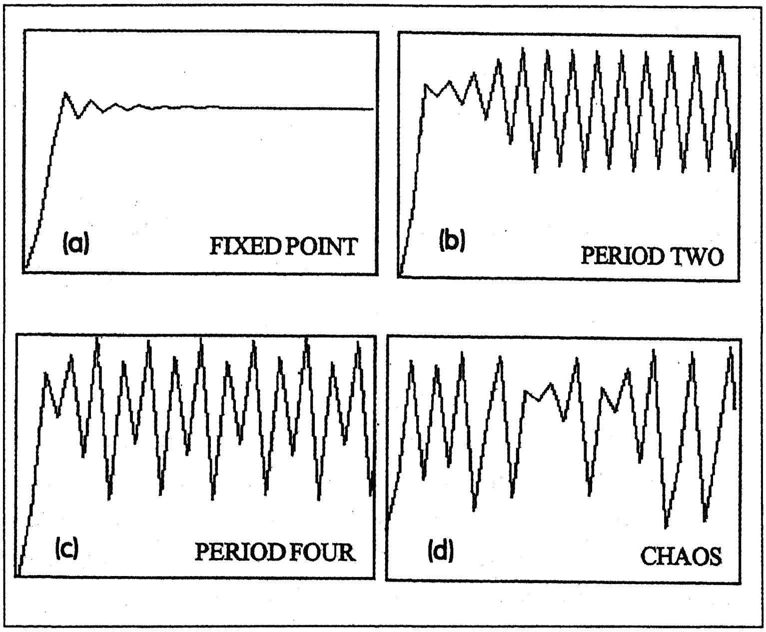
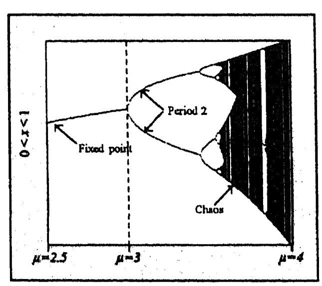


Figure 3 *Behaviour of $x_{n+1} = \mu\, x_n\, (1 - x_n)$ for various μ values. The graph shows values of x_n (along the y axis) plotted against increasing values of x_n in the x axis.* (a) *when μ ranges between 1 and 3.* (b) *when μ ranges between 3 and 3.44 (approx.)* (c) *when μ ranges between 3.44 (approx.) and 3.54 (approx.)* (d) *graph for chaotic regime.*

successive period doublings. Period two gives way to period four, which gives way to period eight and so on. The behaviour of the system as the parameter μ is increased is shown in *Figure 4*, generated by the computer. Along the x-axis, the μ value increases from 2.5 to 4. The y-axis represents the long-term behaviour of the system — i.e. the value that x_n settles down to finally. It can be seen that the bifurcations occur faster and faster and suddenly break off. Beyond 3.57, the periodicity of the system gives way to chaos and long term values of x_n do not settle down at all. In the midst of this complexity, stable cycles of periods such as 3 or 7, suddenly return (called *periodic windows*), only to break off once again into chaos. In fact there is an interesting theorem, which states that in any one-dimensional system, if for some parameter value, the system has a period of three, then the same system (for some other parameters) will also display periods of all other values, as well as completely chaotic behaviour.

Chaos occurs even in the most deceptively simple systems.

Figure 4 *Bifurcation diagram of the equation $x_{n+1} = \mu . x_n\ (1 - x_n)$.*



The period-doubling cascade is one of the standard routes to chaos. It occurs in many non-linear systems that depict chaotic behaviour. For example, in the following dynamical system, period-doubling cascades occur before the system

becomes chaotic.

$$x_{n+1} = \mu \sin(\pi x_n) \tag{3}$$
where $0 < x_n < 1$ and $0 < \mu < 1$.

Is there any similarity between systems that take the period-doubling route to chaos? The principle of universality, discovered by Feigenbaum, states that irrespective of the nature of the system, if it takes the period-doubling route to chaos, then the parameter spacings (the range of parameter values for which the period is the same) occur in geometric progression. Feigenbaum computed the ratio of convergence and the constant named after him, is one of the fundamental physical constants. Later, Feigenbaum's results were reformulated in a rigorous mathematical framework but the initial insights were provided by numerical simulations.

The period-doubling cascade is one of the standard routes to chaos.

Feigenbaum's principle of universality can be stated rigorously as follows. Let μ_0 be the parameter value for which the first period doubling occurs (in other words μ_0 is the parameter value at which period 1 gives way to period 2). Let μ_1, μ_2, μ_3 be the succeeding period doublings. Now we define F_k as,

$$F_k = [\mu_k - \mu_{k-1}] / [\mu_{k+1} - \mu_k] \quad \text{for } k = 2, 3, 4\ldots .$$

The sequence F_2, F_3, F_4 ... converges to a number F_∞ called *Feigenbaum constant.* F_∞ is the same for all systems taking the period doubling route to chaos.

In the next section, we take two dissimilar systems and compute the geometric convergence ratio of the parameter spacings and thereby verify the principle of universality.

Computer Assisted Verification of Universality

An algorithm for computing the parameter value at which a period doubling occurs is presented. The algorithm can be better understood from the accompanying flow chart *(Figure 5)*. The para-meter values at which the period-doublings occur can then be used to compute the parameter spacings and the spacing ratios.

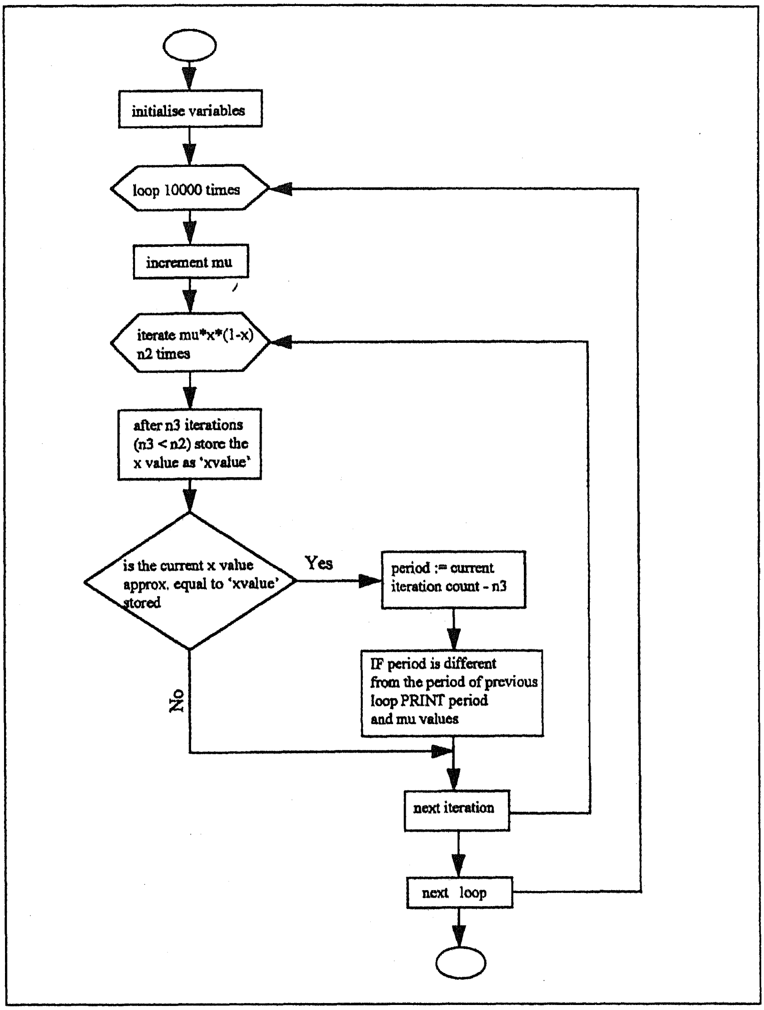


Figure 5 Flowchart for computing the parameter values for which period doubling occurs.

The principle of universality, discovered by Feigenbaum, states that irrespective of the nature of the system, if it takes the period-doubling route to chaos, then the parameter spacings occur in geometric progression.

The idea of the algorithm is to increment the parameter values in an outer loop and to iterate the equation in the inner loop. ABS is a function that returns the absolute value of the input value. RAND is a procedure that returns a random number between 0 and 1. The first few iterations (an arbitrary choice of 8000 iterations in the algorithm) are used to allow the transients to die down. At a fixed *count* value (8000 in the algorithm and n3 in the flow chart) the value of x is stored in the *x value* variable. In the subsequent iterations, we monitor if the value stored in *x value* arrives again. The *count* value for which *x value* arrives again less the *count* value at which *x value* was stored (8000 in this case) gives us the *period* of the system. If the current period (given by *period*)

```
/* Variables used :
        pcount, count        : loop counter variables
        prevperiod, period   : stores the period values
        r                    : stores the parameter values
        x, xvalue            : used to store the actual iteration values
        flag                 : boolean that determines when to exit out of the WHILE loop. */

  prevperiod, pcount, count, period : Integer;
  mu, x, xvalue : float;
  flag : boolean;
  prevperiod := 0;
  FOR pcount := 0 TO 10000          /* outer loop for incrementing parameter value */
      mu := 3.43 + 0.15/10000 * pcount;     /* increment the parameter value */
      count := 1; x := RAND; flag := TRUE;  /* initialise variables before inner loop */
      WHILE    count < 15000 AND flag = TRUE DO
               count := count + 1;
               x := mu * x * (1-x);          /* perform iteration of the system*/.
               IF count = 8000 THEN
                    xvalue := x;              /* store value of x at 8000th iteration */
               END IF
               IF count > 8000 THEN    /* monitor and see when xvalue returns */
                    IF ABS(xvalue - x) < 0.00001 AND x > 0 THEN
                         period := count - 8000;
                    ENDIF
                    IF period <> prevperiod THEN
                         PRINT period, mu;
                         prevPeriod := period;
                    END IF
                    flag := FALSE;
               END IF
      END WHILE
  NEXT pcount
```

The initial discovery of the universal constant by Feigenbaum was done using a pocket calculator.

is different from the previous period (stored in *prevperiod*), the period value (*period*) and the parameter value (r) is printed. The parameter spacing and Feigenbaum constant are then calculated manually.

The algorithm was applied to two dissimilar systems $x_{n+1} = \mu.\ x_n\,(1-x_n)$ and $x_{n+1} = \mu.\sin(\pi.\,x_n)$. The results obtained are in *Tables 1* and *2*.

The numerical studies indicate that for non-linear, one-dimensional systems, which take the period doubling route to chaos, the ratio of the spacing of the parameter values could be universal, in the sense that, for a wide class of systems, it is independent of the details of the equations. The above experiment provides a good metaphor for the role that computers have played in the field of chaos. In fact, the initial discovery of the universal

Table 1. Analysis of the logistic equation $x_{n+1} = \mu . x_n (1-x_n)$

Period#	starting μ value	parameter spacing	Feigenbaum constant
1	1.0000000000	2.0000000000	
2	3.0000000000	0.4494897427	4.44948
4	3.4494897427	0.0946228040	4.75033
8	3.5441125467	0.0203330686	4.65364
16	3.5644456153	0.0043638742	4.65940
32	3.5688094895	0.0009362636	4.66064
64	3.5697457531		

Table 2. Analysis of the sine function $x_{n+1} = \mu \sin (\pi x_n)$

Period#	starting μ value	parameter spacing	Feigenbaum constant
1	0.3000000000	0.4198154390	
2	0.7198154390	0.1134053575	3.70190
4	0.8332207965	0.0253723920	4.46963
8	0.8585931885	0.0054861570	4.62480
16	0.8640793455	0.0011784030	4.65558
32	0.8652577485	0.0002528400	4.66066
64	0.8655105885		

constant by Feigenbaum was done using a pocket calculator, using a procedure similar to the one used above.

Conclusion

The discovery of chaos has far reaching implications in many branches of science. Chaos has provided us with a new way of looking at nature. It has helped us to find order in places where we earlier found only disorder. In conclusion, two points are worth reiterating. First, chaos can occur in deceptively simple dynamical systems. Second, the computer is an essential and indispensable tool in studying chaotic dynamics.

Suggested Reading

Robert May. Simple Mathematical Models with Complicated Dynamics. *Nature*. 261:985-992. 1976.

Benoit B Mandelbrot. The Fractal Geometry of Nature. Freeman, San Francisco. 1982.

H O Peitgen, P Ritcher. The Beauty of Fractals. Springer-Verlag, New York. 1986.

James Gleick. Chaos : Making a New Science. Viking, New York. 1987.

Denny Gulick. Encounters with Chaos. McGraw Hill, New York. 1992.

Ian Stewart. The Problems of Mathematics. Oxford. 1992.

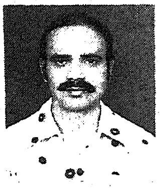
Ian Stewart. *The Mathematical Intelligencer*. 17:52. 1995.

Address for correspondence
Balakrishnan Ramasamy
Siemens Communication Software, 10th floor
Raheja Towers
26-27, M G Road
Bangalore 560 001, India.

Echolocation

The Strange Ways of Bats

G Marimuthu



G Marimuthu has studied the behaviour of bats for almost two decades. His pioneering experiments have led to an understanding of how bats catch frogs in total darkness.

Bats are capable of avoiding obstacles that they encounter, even in complete darkness. This is because they emit ultrasound (high frequency sound) and analyse the echo produced when the sound hits objects on their path. This article describes the hunting flight of bats and how echolocation is useful in prey capture. Prey capture without the aid of echolocation by some bats is also described.

The March 1996 issue of *Resonance* introduced us to the fascinating world of bats, the only flying mammals of the world. As opposed to traditional views on bats, they are a harmless and interesting group of animals. Awareness about bats and the need to conserve them has increased considerably in recent years. A very interesting feature which was only briefly mentioned in *Resonance* Vol.1, No.3, is the ability of bats to 'see' through their ears. The microchiropteran bats use a special property called 'echolocation', both to avoid obstacles on their way and to locate and capture their prey.

Echolocation is a specialized process of orientation used by bats. Bats emit high frequency sound waves while navigating, and process the echo that comes back from obstacles. This method assists prey location and capture.

Bats emit high frequency sound waves while navigating, and process the echo that comes back from obstacles. This method assists prey location and capture.

The Discovery of Echolocation

During the year 1790, Lazzaro Spallanzani, an Italian naturalist, first observed that bats were able to avoid obstacles while flying even in total darkness. He also found that despite the surgical removal of eyes, bats could fly without bumping into obstacles. Later, Charles Jurine, a Swiss zoologist, plugged the ears of bats

and observed their inability to perform these correct orientations. Spallanzani repeated these experiments and obtained similar results. Both of them concluded that bats could 'see' through their ears! The French naturalist Cuvier disagreed with this statement. He explained that a sense of touch in the wing membrane caused the bats to avoid obstacles. In 1920, Hartridge, a British physiologist put forward the hypothesis that bats emit ultrasound and listen to the echoes of these sounds. After 18 years, the American zoologist, Donald R Griffin along with Pierce, a physicist, used a microphone sensitive to ultrasound and demonstrated that bats do emit trains of ultrasonic pulses while flying. They showed that the number of sound pulses increased as bats approached obstacles on their flight path. They also noticed that the bat's mouth was always open when the sounds were emitted. Griffin continued the experiments and found that closing the mouth of the bat resulted in disorientation. He established that bats emit sounds through their mouths. It was Griffin who coined the term 'echolocation' in 1938. In 1958, he published his classic book, 'Listening in the Dark' which documents many details about the discovery of echolocation. Echolocation is one of the methods of orientation mainly used by the microchiropteran or insectivorous bats. While flying, these bats emit high frequency ultrasound. These sound pulses hit obstacles like rocks, trees, walls etc. and their echoes are heard by bats. By analysing these echoes, bats are able to find their way even deep into underground caves in which there is absolutely no light.

Donald R Griffin showed that bats emit ultrasonic pulses.

Vocalizations of Bats

Like other mammals, including humans, bats emit sound through the voice box or larynx. Sound is produced when the vocal chords vibrate as air passes over them. Hence these sounds are called vocalizations. The muscles in the larynx adjust the tension on the vocal chords. This controls the rate of vibration of the vocal chords which explains the frequency or pitch of the sound emitted. Some of the characteristics of sound are shown in the

It was Griffin who coined the term 'echolocation' in 1938. In 1958, he published his classic book, 'Listening in the Dark' which documents many details about the discovery of echolocation.

Figure 1 The face of the Indian false vampire bat Megaderma lyra. It is a microchiropteran and carnivorous bat. It weighs about 40 g. While flying, it emits ultrasounds through its nostrils, which help to beam the sound pulses. The huge pinnae are able to collect the faint noise created while the prey moves.

box. Most of the species (eg. Indian pygmy bat, free-tailed bat, tomb bat) emit their echolocation sounds through the mouth. A few other species (eg. Indian false vampire bat, leaf nosed-bat, horseshoe bat) produce their vocalizations through the nostrils. The latter species have grotesque facial ornamentations. This is known as the noseleaf (*Figure 1*). It is a shallow, parabolic portion surrounding the nostrils and a spear shaped, rounded or fleshy superior portion. The structure of the noseleaf varies from species to species. The noseleaf serves to narrow and focus the outgoing beam of sound.

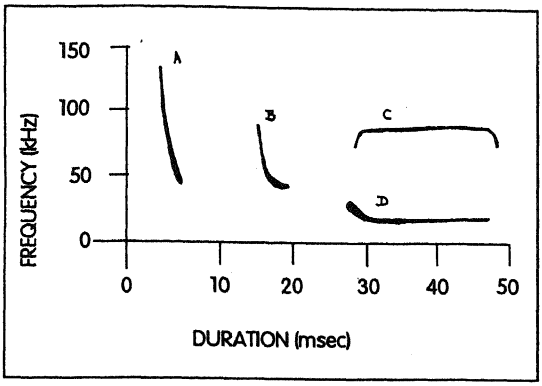
Vocalizations used in echolocation are generally divided into two categories.

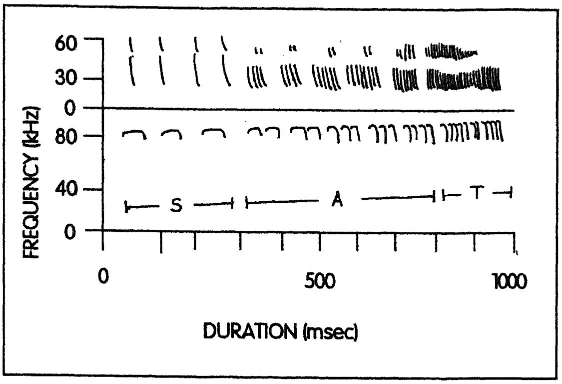
- *Broadband signals* These cover a wide range of frequencies, from 20 to 140 kHz and have shorter durations of less than 5 milliseconds. They are technically called frequency modulated (FM) pulses. Each pulse starts at a high frequency and sweeps down to lower frequency within a short duration.
- *Narrowband signals* These have a constant frequency (CF) and longer durations of about 100 milliseconds.

Functions of Echolocation

The hunting flight of bats is divided into three stages: the search stage, the approach stage and the terminal stage.

Even though there are two such distinct kinds of sounds (*Figure 2*), bats use either one or combinations of both depending on the situation and gather detailed information on their flight path. The hunting flight of bats is divided into three stages: the search stage, the approach stage and the terminal stage (*Figure 3*). During the search stage, bats emit sound pulses with a low repetition rate of about 10 pulses per second. Actually a correlation exists between the habitat in which a species regularly forages and the type of signal it emits at this stage. Usually short CF pulses, with or without an FM tail, are found in species that forage in open spaces where vegetation and other obstacles are not found. Bats that hunt close to vegetation or the ground, emit pulses which mainly have an FM sweep. Theoretically, the





amount of information available from a signal is proportional to its bandwidth. A broadband outgoing sound pulse would cause a greater number of altered frequencies in the returning echoes. Bats use such echoes to analyse the features of the target, for example to differentiate prey from the background clutter and to differentiate smooth and rough surfaces suitable for landing. They can accurately discriminate between targets that are within 10-15 mm of each other. To estimate the target range (distance), bats analyse the time delay between the emission of sound and its return as echo just like a radar detects objects several metres away.

***Figure 2** Sonograms of different types of echolocation sounds shown as frequency in the ordinate and duration in the abscissa scales : (A). Steep broadband FM signal starts at a higher frequency and ends in a lower frequency in a short duration. (B). Steep FM signal ends with a shallow FM component. (C). A long CF narrowband signal with an initial increasing FM and a decreasing FM tail at the end. (D). Signal starts as a shallow FM with a long CF component.*

***Figure 3** Pattern of the emission of echolocation pulses by bats which emit only FM signals (top) and bats which emit a combination of both CF and FM signals (bottom) at three different stages. (S) - search stage, (A) - approach stage, (T) - terminal stage.*

A few other species like horseshoe bats which emit narrowband signals with longer CF component use an alternative strategy of echolocation. They distinguish the moving prey from nonmoving obstacles by means of the Doppler effect (see box and *Resonance*, Vol.1, No.2, Page 14).

The main function of the search stage is to detect the potential prey among obstacles. The big brown bat *Eptesicus fuscus* can detect a sphere having a diameter of 2 cm, at a distance of 5m. The same bat detects a 0.5 cm sphere at a distance of 3 m.

The onset of the approach stage represents the first visible reaction of the bat to the target. This stage begins when the bat is between 1 or 2 metres away from the target. The bat turns its head and ears towards the target. It also increases the repetition rate of the echolocation sounds to about 40 pulses per second. In

Characteristics of Sound

Sound is a series of vibrations in air or water for example, picked up by the ears and interpreted as a sensation by the brain. A few characteristics of sound are relevant to echolocation. *The frequency or pitch* of the sound of bats is measured in kilohertz, abbreviated as kHz. One kHz is one thousand cycles per second or 1000 Hertz. Humans can hear up to 20 kHz. Sounds having a higher frequency than this are called ultrasound. The echolocation calls of bats are inaudible to humans and hence called *ultrasonic*. Since high frequency sounds are more rapidly absorbed by the atmosphere, the echolocatory system works within a limited distance. The *intensity* of sound is measured in decibels, abbreviated as dB. This unit is related to the ratio of the sound intensity to a standard, which is taken to be the threshold sound intensity detectable by the human ear. *Table 1* provides the decibel scale to measure the loudness of various sounds. The echolocation sound of bats is about 110 dB at 10 cm in front of a bat's mouth. This is slightly more intense than the sound from a milk cooker, a common vessel in the kitchen.

Theoretically a bat receives the echo of its sounds within 500 milliseconds (1000 milliseconds = 1 sec). The obstacles could be away at a maximum distance of about 85 m. The bat detects the distance of the target by measuring the time interval between the emitted signal and its echo. The range of frequencies of the echolocation pulses is the bandwidth of the signal. The power spectrum explains the distribution of energy on the frequencies of the signal (see *Figure 4*). A bat collects detailed information about targets by comparing the power spectra of the emitted sound and its echo.

Table 1. The decibel scale

dB	*Examples*	*dB*	*Examples*
10	Rustling leaves	80	Vacuum cleaner
20	Whisper	90	Classroom in a school
30	-	100	-
40	Voices in city night	110	A road drill
50	Normal speech	120	-
60	A busy super market	130	Jet aircraft take off
70	-	140	Painful sounds

The hearing sensitivity of bats is much higher than other mammals.

addition to these changes, a qualitative change in the pulse pattern also occurs. In species (eg. *Myotis myotis*) which emit only FM pulses, the slope of the FM sweep becomes steeper, the duration of the pulse becomes shorter but the bandwidth of the

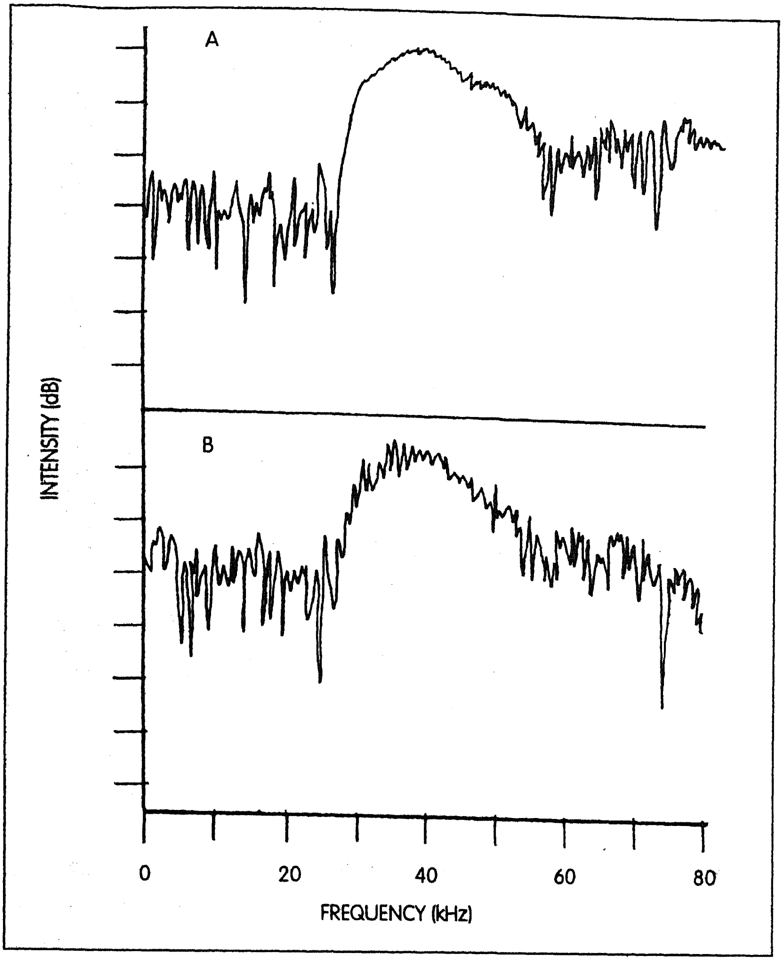


***Figure 4** The spectrogram of the echolocation call (A) and its echo (B). The spectral difference between the pulse and the echo provides detailed information about the target structure.*

signal remains the same. In a few other species like *Nyctalus noctula* which emit only CF pulses during the search stage, an abrupt switch to emitting brief FM pulses occurs. The CF component is dropped. Horseshoe bats which use long CF-FM pulses during the search stage do not drop the CF component at the approach stage. Their pulses become shorter with an increase in bandwidth of the FM component.

The big brown bat *Eptesicus fuscus* can detect a sphere having a diameter of 2 cm, at a distance of 5m. The same bat detects a 0.5 cm sphere at a distance of 3 m.

Thus a shift towards emission of FM pulses is discernible at the approach stage. Since the information content is greater in the broadcast (FM) signals, this shift is useful to decide whether to catch a prey or to avoid an obstacle or to land on a roosting site.

Doppler Shift

Our ears hear a changed sound when we listen to a sound source which moves rapidly towards or away from us, eg. a car passing us with its horn blowing. We experience a sudden drop in frequency as the car passes away from us. Even though we hear a sudden change in frequency, the horn actually sends out sound waves at a regular interval. If we stand ahead of the car (person 'A' in *Figure 5*) our ears receive more than the normal number of sound waves and we hear a higher frequency than the real tone of the horn. After the car passes, our ears receive fewer sound waves (person 'B' in *Figure 5*) so that the frequency becomes lower, with a sudden drop at the moment the car passes us. The faster the car moves, the greater the change in frequency. This effect of motion on the frequency of sounds was first pointed out by an Austrian scientist Christian Doppler and it is named after him as *Doppler shift.*

The long constant frequency signal of the echolocation sounds emitted by a few species of bats is used for measuring the Doppler shift but is not suitable for target description. Bats are able to analyse the shifts that occur in the echo frequency produced by a flying insect. They use this method to detect the insect prey from the large amount of echo clutter produced by the dense foliage or other background objects.

When bats reach the target within a distance of 50 cm, the terminal stage begins. A steep increase in the repetition of the emission of about 100 or even 200 pulses per second occurs. This increased rate rapidly updates the information and the bat makes the final decision whether or not to catch the prey. This rapid increase in the emission of sound pulses during the terminal stage is termed as 'final buzz'. In most bat species, the sound pulses emitted at this stage are only FM sweeps with three or four harmonics. These are of lower duration of within 0.5 milliseconds. In bats which use the Doppler shift, like the horseshoe bats, the CF component still remains but is reduced in duration and is about 10 milliseconds (compared to 60 milliseconds during search stage). After detecting the insects, bats capture them by using their wing membranes and transfer them to their mouth.

A few other species like horseshoe bats distinguish the moving prey from nonmoving obstacles by means of the Doppler effect.

The hearing sensitivity of bats is much higher than other mammals. This specialization allows bats to receive and analyse faint echoes. When the echoes return to them, they are received by the auditory system similar to other mammalian patterns.

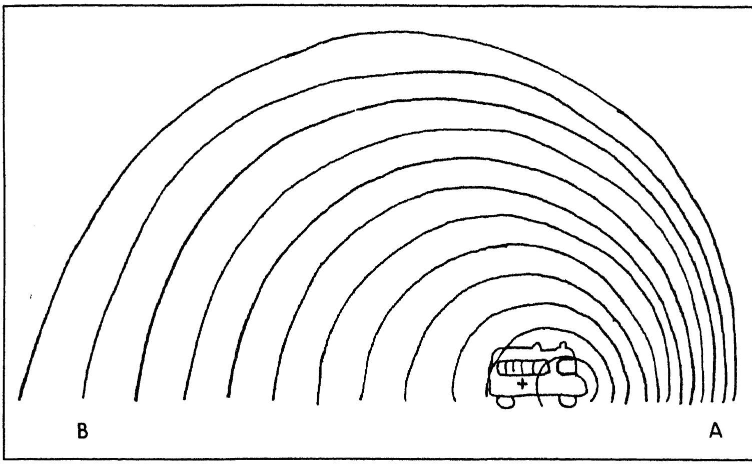


Figure 5 Each sound wave starts out as a circle. Since the horn is moving forward continuously, the centre of each circle is a little farther along the road than the previous one. This makes the wave 'crowded' (high frequency) in front (A) and 'stretched out' (low frequency) at the back (B).

Echolocating bats have prominent external ears. Their pinnae are specialized to amplify the faint echoes. The mechanical vibrations of the echoes travel through the ear drum, middle ear and reach the cochlea of the inner ear. A helical ribbon, known as the basilar membrane, present in the cochlea contains hair cells. These are the receptor cells that convert the mechanical vibrations of the echoes into electrical signals and transmit them to the brain along the auditory nerve. Processing of the echoes takes place in the brain. The processing includes information such as the pulse-echo delay and comparison of spectral features of the original sound and its echoes. From this process a bat gets an 'acoustic picture' of its flight path.

Prey Capture without Echolocation

Recent studies show that some species of bats do not use echolocation to detect their prey. These are the false vampire bats in India, Australia and Africa, long eared bats in North America, mouse eared bats in Europe, fringe lipped bats in Panama and slit-faced bats in Africa. The Indian false vampire bats listen passively to the noise associated with the movement of the prey (frogs, mice, larger insects, etc.). The fringe-lipped bats use the songs of male frogs to locate and capture them.

Recent studies show that some species of bats do not use echolocation to detect their prey.

They can distinguish the calls of edible frogs from those of

poisonous toads. The vampire bats (living only in Central and South America) use the breathing noise of the cattle to locate and to feed upon their blood. All these species of bats produce faint echolocation signals but use them only to gather information about the background. They are hence known as whispering bats. Echolocation is a unique and fascinating characteristic of bats.

The vampire bats of Central and South America use the breathing noise of the cattle to locate and to feed upon their blood.

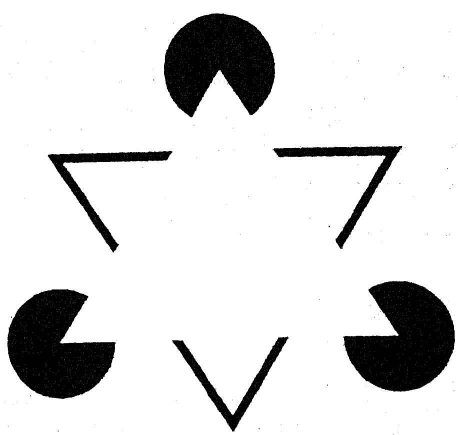
Address for correspondence
G Marimuthu
Department of Animal Behaviour & Physiology,
School of Biological Sciences,
Madurai Kamaraj University,
Madurai 625 021. India.

Suggested Reading

D R Griffin. Listening in the Dark. Yale University Press, New Haven. 1958.

G Neuweiler. Echolocation and Adaptivity to Ecological Constraints. In: Neuroethology and Behavioural Physiology. (Eds) F Huber and H Markl. Springer Verlag, Berlin. 1983. 280-302.

D Young. Nerve Cells and Animal Behaviour. Cambridge University Press. Cambridge. New York, Sydney. 1989.



Kanizsa triangle ... consists of illusory contours. A normal visual cortex sees a triangle even though interconnecting lines are missing. Such illusions show that the visual cortex must resolve conflicts between different functional areas.

Honeybees can see optical illusions ... Brazilian researchers have found that bees rewarded with a sugar solution can be taught to "see" Kanizsa triangles.

Sampling, Probability Models and Statistical Reasoning

Statistical Inference

Mohan Delampady and V R Padmawar

Statistical inference is introduced here as an application of inductive inference. Using examples it is illustrated how random sampling allows data to be modelled with the help of probability models and how these probability models provide the mathematical tools for statistical inference.

Mohan Delampady received his Ph.D. from Purdue University in 1986. After spending five years at the University of British Columbia he joined Indian Statistical Institute in 1991 and has been with its Bangalore centre since then. His research interests include robustness, nonparametric inference and computing in Bayesian statistics.

V R Padmawar is with Indian Statistical Institute, Bangalore. He received his Ph.D. from Indian Statistical Institute in 1984. His research interests lie in the area of sampling theory.

Inductive Inference

Experimentation is a vital ingredient of scientific advancement. We have all conducted experiments in school or college laboratories at one time or another. For example, consider the simple experiment involving a pendulum to determine the gravitational constant. We obtain data on l, the length of the pendulum and t, the time required for some fixed number, say m, of oscillations. We then substitute these values in the following formula to get the value of the gravitational constant g.

$$g = 4\pi^2 \times \frac{l}{(t/m)^2}$$

We repeat the experiment a few times and for each trial compute the value of g. We then compute the average of these values. Incidentally, most instructors would also want us to report the *experimental error*. We compute the standard error of the values of g based on different trials to get an estimate of the *experimental error*. (The meaning of these *error* terms will become clear as you go on!)

Having conducted the experiment we are willing to accept the

value of g that we obtained to hold even outside our specific location. In other words, for example, if the experiment were conducted in Hassan rather than in Bangalore we believe that we would still get more or less the same value of g. This seems to be acceptable to us without actually conducting the experiment in Hassan. Scientists often generalise from a particular experiment to a class of 'similar' experiments. This kind of extension from the particular to the general is called *inductive inference*. This is in contrast to *deductive inference*, an example of which is a mathematical proof of a conjecture. Clearly, inductive inference entails a certain degree of uncertainty.

Scientists often generalise from a particular experiment to a class of 'similar' experiments. This kind of extension from the particular to the general is called *inductive inference*.

One need not be disheartened by this element of uncertainty if the degree of uncertainty can be estimated. This can indeed be done provided we follow certain principles. Statistics plays an important role in providing techniques for making an inductive inference as well as for measuring the degree of uncertainty in such an inference. This uncertainty is measured in terms of probability.

Let us now consider an example of inductive inference. A consignment of 10,000 units each having its own identification number comes to the office of a statistical officer in a firm. A unit can be classified as defective or nondefective depending on certain specified standards and criteria. The officer's job is to decide whether to accept or reject the consignment in its totality. As common sense dictates, we would accept the consignment if there weren't 'too many' defective units, or in other words, if the proportion of defective units in the consignment is not very substantial. Let θ denote this unknown proportion of defectives. Our objective now is to determine this θ. One way of achieving this is to test each and every unit by subjecting it to the specified standards and criteria. This would tell us whether a given unit is defective or not. This in turn would give us the total number or equivalently the proportion of defective units in the consignment. Such a procedure is often long drawn and expensive. In

many situations such as measuring the breaking strength, we end up destroying the units while testing them. Isn't it futile if at the end of the procedure, we know the exact proportion of defectives θ in the consignment but are left with no usable units?

A thought that comes to mind is whether one can subject only a few units to the specified standards and criteria and based on these few units make a statement about the unknown θ. The answer is that we cannot determine the exact value of θ but can make a probabilistic statement about it provided we select the few units in accordance with certain principles. 'Random sampling' which is discussed later in some examples is one such technique to select a few units from the entire 'population' that satisfies the above requirement. It is shown later that random sampling allows the data obtained to be modelled using probability models. These probability models in turn provide the mathematical tools for statistical inference.

In science as well as in human affairs we often lack resources to study more than a glimpse of the phenomenon that might advance our knowledge!

Moreover, if we adopt the above testing procedure only for a sample from the consignment rather than the whole consignment, we would save on cost, time, as well as effort. Prior to the advent of pressure cookers many of us have seen how our grandmothers would mash a few grains of rice to determine whether the entire rice in the vessel was properly cooked or not. Thanks to their wisdom we were never left starving and we seldom ate undercooked rice. In science as well as in human affairs we lack resources to study more than a glimpse of the phenomenon that might advance our knowledge!

Thus, we select a sample of a few units from the consignment, observe the number of defective units in the sample, and from this knowledge try to predict the unknown proportion of defectives θ in the consignment. We cannot be certain of our answer but we can make a statement about the error in our answer. This gives us an idea about inductive inference.

Even in our problem we could have asked ourselves the following two questions :

1. What is the value of θ?
2. Is $\theta \leq .01$(say)?

The first question leads to the *theory of estimation* whereas the second question leads to the *theory of testing of hypotheses*. In what follows we discuss some simple examples where statistical inference can be done meaningfully.

Probability Models and Statistical Inference

A random sample is one where every unit in the population has the same chance of being included in the sample.

Example 1. Suppose that some of the printed circuit boards manufactured by a company have a certain defect which can only be detected with extensive testing. A "random sample" of n of these boards is chosen for testing. A random sample is one where every unit in the population has the same chance of being included in the sample. Suppose k out of these n turn out to be defective and the rest are fine. What can be said about the proportion θ of defective boards in the entire population of boards that this company has produced, if we can assume that the total number of boards produced is very large?

First, note that θ is also the probability that a randomly chosen board (from the population) is defective. Now, let X denote the number of defective boards in the sample of n boards which were selected. For any integer x between 0 and n, we claim that

$$P(X = x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \tag{1}$$

This can be justified as follows. Sampling n units from a population involves n trials of the same experiment, i.e. that of choosing a unit at random. Since the population size is huge, we can assume that the outcomes of different trials are independent. Therefore the probability of obtaining any given sequence of outcomes

which consists of x defectives and $n-x$ non-defectives is $\theta^x (1-\theta)^{n-x}$. Since there are $\binom{n}{x}$ such sequences, the probability of the event $\{X = x\}$ must be what we specified above. (Refer to Karandikar, R L 1996, On Randomness and Probability, ***Resonance***, Vol. 1, No. 2, pp. 55-68 for related material.) The probability distribution specified by (1) is called the *binomial distribution*. Using (1) it can be shown easily that the expectation of X is $E(X) = n\theta$ and the its variance is $Var(X) = n\theta(1-\theta)$.

For each θ in the unit interval (0, 1), (1) gives a different binomial probability model for X. What is the model in which observing the event $\{X=k\}$ is most likely? Let us denote by $P(X = x | \theta)$ the probability specified by (1) for the given θ. Let us define the function $L(\theta)$ by

$$L(\theta) = P(X = k/\theta),$$

where k is the observed number of defectives (which is known now). Note that L is a nonnegative function defined on the set of all possible values of θ which is called the parameter space. This function L is called the likelihood function of the unknown quantity θ and clearly this measures how likely is the event $X=k$, if θ is indeed the true value of the proportion of defectives in the entire population.

One appealing method to *estimate* θ is to find the value of θ which maximizes the likelihood function. This is called maximum likelihood estimation and the interpretation of the obtained estimate is that it gives the model, from all the models considered in (1) in which the observed event is most likely.

In the problem above, the maximum likelihood estimate of θ is $\hat{\theta} = k/n$ since this is the value of θ which maximizes $\binom{n}{x}\theta^k (1-\theta)^{n-k}$. Since X denotes the (random) number of defectives in the sample of (fixed) size n, we see that the maximum likelihood estimation method provides X/n as the estimator for θ. Using the *Weak Law of Large Numbers* (see Karandikar, 1996)

Polya's View

"Experience modifies human beliefs. We learn from experience or rather, we ought to learn from experience. To make the best possible use of experience is one of the great human tasks and to work for this task is the proper vocation of scientists. A scientist deserving this name endeavours to extract the most correct belief from a given experience and to gather the most appropriate experience in order to establish the correct belief regarding the correct question."

it can be shown that the maximum likelihood estimator, X/n approaches θ as $n \rightarrow \infty$. It is comforting to note that if a large number of boards chosen at random are tested it is indeed possible by the above method to determine the actual proportion of defectives.

The precision of an estimator can be measured by its *standard deviation* (again refer to Karandikar, 1996 for details) if the estimator's expectation is the quantity it is supposed to estimate. Note that, smaller the standard deviation better the estimate on average. In the present case, $E(X/n) = n\,\theta\,/n = \theta$ and $Var(X/n) = n\,\theta\,(1-\theta)\,/n^2 = \theta\,(1-\theta)/n$. Therefore the standard deviation of the estimator X/n is $\sqrt{\theta(1-\theta)/\,n}$. For the observed data, namely k defectives in the sample of n, θ is being estimated by $\hat{\theta} = k/n$ and hence a measure of precision of the estimate is given by $[\,\hat{\theta}\,(1-\hat{\theta}\,)/\,n\}]^{1/2}$.

It is comforting to note that if a large number of printed circuit boards chosen at random are tested it is indeed possible by the maximum likelihood method to determine the actual proportion of defectives.

Example 2. What if θ is very small in the above problem? If n is not very large, most of the samples may show 0 defectives. Then $\hat{\theta}$ is also 0. This is not good enough if we want a good idea about the proportion of defectives in the population. It is clear that the information provided by our experiment is inadequate. How can we modify our experiment to gather more informative data?

One suggestion is to continue (random) sampling until a prefixed number, say r, of defectives is observed. What is the data now? It is simply X = number of sampled boards which obtains r defectives in the sample. It can be seen, arguing as in the above example, that

$$P(X = x) = \binom{x-1}{r-1}(1-\theta)^{x-r}\,\theta^r$$

for any integer $x \geq r$. This probability distribution is known as the *Negative Binomial* distribution. When $r = 1$, it is generally called the *Geometric distribution.* Estimation of θ can be done

following the steps of the above example.

Example 3. Suppose we want to estimate the number of fish in a lake or in a particular part of a sea. Let N denote this unknown number of units. Consider the following experiment which is known as the *capture-recapture* method. Catch N_1 fish from this population. Tag all of them and release them. Now fish again. This time suppose that n fish are caught, n_1 of which have tags on them. Clearly $N_1 \leq N$ and $n_1 < n$. What is an estimate of N and how good is this estimate? N clearly cannot be less than $N_1 + (n-n_1)$ which is the total number of fish that we have seen in the two stages of fishing. For any value of $N \geq N_1 + (n - n_1)$ what is the probability of observing n_1 tagged fish among n fish caught in the second stage? This is exactly equal to the ratio of *the total number of ways of choosing n units from N units such that n_1 of them came from a fixed set of N_1 to the total number of ways of choosing n units from N units.* Therefore, the probability is $p(N) = \binom{N_1}{n_1}\binom{N-N_1}{n-n_1}/\binom{N}{n}$. The estimation procedure discussed above can be easily implemented here too by first obtaining a likelihood function for the unknown parameter N using the above probability expression. As is discussed in Feller (1993), to find the maximum likelihood estimate of N consider the ratio

The *capture-recapture* method is an ingenious way to estimate the number of fish in a lake or in a particular part of a sea.

$$\frac{p(N)}{p(N-1)} = \frac{(N-N_1)(N-n)}{(N-N_1-n+n_1)N}.$$

Note that this ratio is greater than or smaller than 1, according as $Nn_1 < N_1 n$ or $Nn_1 > N_1 n$. Therefore, with increasing N the sequence $p(N)$ first increases and then decreases; it reaches its maximum when N is the integer part of $N_1 n/n_1$, so that the maximum likelihood estimate of N is approximately $N_1 n/n_1$. This is an intuitive estimate since it implies that approximately n/N equals n_1/N_1; it is the same as saying that the observed proportion of tagged fish should be approximately the same as the proportion of sampled fish.

The next example is about elections. Since the elections are just around the corner and various commercial agencies will be making *expert predictions* about which party will win, let us consider the statistical methods they employ!

Example 4. Suppose that we want to know the proportion of eligible voters who support a particular political party. A random sample of size n is selected from this population and suppose k voters support this party. What is a good estimate of the required proportion? How do we obtain a probability model for the experiment just conducted? Let us examine the following simple experiment. Consider a box containing a large number of marbles, a proportion θ of which are red. If n marbles are chosen at random from this box, what is the probability of getting k red marbles? The answer depends on whether the sampling is done with or without replacement. Random sampling with replacement means that the unit sampled is replaced in the box before the next draw is made. In random sampling without replacement the sampled units are not replaced. If the sampling is with replacement the probability distribution of the number (X) of red marbles in the sample is given by the Binomial probability model of (1). If it is without replacement then we get

If the sample size n is very small compared to the population size, the hypergeometric model can be approximated by the binomial model.

$$P(X=x) = \frac{\binom{N\theta}{x}\binom{N(1-\theta)}{n-x}}{\binom{N}{n}}$$

This is the *Hypergeometric probability distribution.* What can be seen is that if the composition of the box doesn't change from draw to draw then the draws are independent and hence we get the Binomial model instead of the Hypergeometric. Note that, if the number of sampled units n is very small compared to the total number N of marbles in the box then the composition of the box doesn't change much from draw to draw even if the sampling is done without replacement. Therefore the Binomial model should be available as the approximation of the Hypergeometric

model for large N. This is indeed true as the following result shows.

$$\lim_{N \to \infty} \frac{\binom{N\theta}{x} \binom{N(1-\theta)}{n-x}}{\binom{N}{n}} = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

To prove this, expand all three terms in the left hand side using factorials, factor out the right hand side from there and then note that whatever remains converges to 1 as $N \to \infty$.

Let us return to the question of predicting the outcome of elections. If a random sample of n voters contains k who support a particular political party, then from the exact Hypergeometric model or the approximate Binomial model, one can check that the maximum likelihood estimate of the proportion of voters in the population who support that political party is simply k/n. Of course, the commercial agencies which conduct the surveys may employ more advanced techniques such as dividing the sample size between different cities according to their population and other important factors. When random sampling involves such additional constraints, the probability models for the sampling design need to be modified accordingly.

If a random sample of n voters contains k who support a particular political party, one can show that the maximum likelihood estimate of the proportion of voters in the population who support the political party is simply k/n.

Example 5. Let us see how our school experiment to determine g can be put in the set up of statistical inference. Let y_i be the value of the gravitational constant that we obtain in the ith trial (for $i=1, \ldots, n$). Then we can represent or model our data as follows:

$$y_i = g + \varepsilon_i, \quad i = 1, \ldots, n,$$

where g is the true (but unknown) value of the gravitational constant and ε_i is the combined error due to various factors which are beyond our control in the ith trial of the experiment. If we assume that there are no systematic errors in the experiment then the expectation of ε is $E(\varepsilon) = 0$. The variance of ε is $Var(\varepsilon) =$

An important point to note is that a probability distribution is used in each of the examples to model data. This is the only way to obtain optimal statistical procedures.

σ^2 which indicates how precise the experiment is. The Gaussian distribution (bell-shaped curve, to be discussed in detail in another article) is normally used to model measurement errors like the ε above. This will imply that our data, $y_1, y_2, \ldots, y_n$ are independent and identically distributed Gaussian (or normal) random variables with expectation g and variance σ^2. It can be shown using the probability density of the Gaussian distribution that the maximum likelihood estimate of g is $\bar{y}$, the average of the observations. This is indeed what we were told to report as the value of g by our instructor, isn't it? To obtain a measure of precision of this estimate, we note that $E(\bar{y}) = g$ and $Var(\bar{y}, y) = \sigma^2/n$. It can be shown again using the probability density of the Gaussian distribution that the maximum likelihood estimate of σ^2 is $\hat{\sigma}^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2/n$. Therefore, an estimate of the experimental error is $\hat{\sigma}/\sqrt{n}$.

In the above discussion we tried to give a flavour of statistical inference using some simple illustrations. The important point to note is that a probability distribution is used in each of the examples to model data. This is the only way to obtain optimal statistical procedures. It can be readily seen that the scope of statistics is much wider than what is discussed above using mostly simple discrete probability models to estimate unknown parameters. One very important topic which is not covered at all is hypothesis testing. Likelihood methods again play a substantial role here as well. The methodology involved here is material for a future article.

Suggested Reading

G Polya. Induction and Anology in Mathematics. Princeton Univ. Press.1954.

G Polya. Patterns of Plausible Inference. (2nd ed.). Princeton University Press. 1968.

M A Mood, F A Graybill, and D C Boes. Introduction to the Theory of Statistics. McGraw-Hill Kogakusha. (3rd Ed. Intl. Student Edition), 1974.

W Feller. An Introduction to Probability Theory and Its Applications. Wiley- Eastern, New Delhi. Vol.1 (Third Edition).1993.

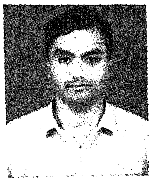
Address for correspondence
Mohan Delampady
Indian Statistical Institute,
8th Mile, Mysore Road
Bangalore 560 059, India.

What's New in Computers

The Java Internet Programming Language

T S Mohan

The Java programming language is poised to change the way Internet is currently being used. In this article, we explore Java and find out why it has become popular with programmers all over the world.



T S Mohan has been working on distributed programming languages and environments for many years. Of late he has been hacking with Java and HotJava. When he doesn't hack, he hitch-hikes around Bangalore. He is currently a Senior Scientific Officer in the Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore.

Introduction

Computers and telecommunications have enabled our progress into the *information society*. The Internet serves as the *information superhighway* that every nation can connect to. It essentially consists of powerful computers with a variety of capabilities, resources and information networked via high-speed and large bandwidth communication channels. These computers are also known as *servers*. Connected to these servers are a large number of desktop systems using which many computer literate participants interact. They can search and obtain the latest information on any topic. In addition, they can share their unique experiences and information, by adding them to the Internet. Information over the Internet is going multimedia; that is, it will include text, graphics, images and sound.

The Internet is a rapidly evolving network of computers, where chunks of related information are available across the world. This information is always evolving, thus making all stored versions obsolete within a short period. It was a challenge to create information structures that looked like a simple document to a user on a computer system, but were actually spread across multiple systems in different parts of the world. This distributed document enabled authors to build-in references to remote documents. The creators of these remote documents could always update them whenever necessary and these updates were always available at any access of the original document.

The Internet is a rapidly evolving network of computers, where chunks of related information are available across the world.

The Global Information Society

When the comet Shoemaker-Levy broke up into pieces and crashed into Jupiter, the images sent back by the spacecraft Galileo were immediately made available over the Internet by NASA. Unable to meet the access demand by users – astronomers, scientists and lay public – all over the world within hours of release, NASA scientists had to replicate this information in other systems that were geographically spread out. This helped to bring the data transfer load on the computer networks under control. Thus people all over the world had direct and immediate access to the images sent by Galileo even as the NASA scientists themselves were using it as part of their scientific study. Never before had this kind of information been distributed or used and the results of the efforts by different people studying the rare astronomical phenomenon been shared so easily. The global information society has arrived. The network is the computer.

Initially developed at the CERN labs in Europe and distributed freely over the Internet, the first *world wide web browser* was essentially a *hypertext* document interpreter package that helped follow hypertext links or references to other documents residing in different sites. With this came the distributed hypertext model of transparently structuring information on computer systems, along with a new information transfer protocol for the Internet called *hypertext transmission protocol* or http. A language was developed to express documents in the hypertext mode. This was called the *hypertext markup language* or html. People quickly realized that sending html based documents over the Internet was the ideal way to publish and annotate scientific research and reference material.

People quickly realized that sending html based documents over the Internet was the ideal way to publish and annotate scientific research and reference material.

Augmenting the web browsers with windows based graphical user interface (as in Mosaic and Netscape) made it popular. Currently, hypertext documents comprise text, graphics, sound and images. When sophisticated web documents with features that used non-standard extensions to the html were made, most web browsers around the world ignored them because the underlying computer systems were incapable of supporting them. Examples include documents with very advanced graphics made using graphics workstations supporting more than a million

color palettes, video clippings made with video capture and display package, as well as digitized sound made using audio recording and playing package. In addition, the transmission of such documents over the computer network hogged a large bandwidth. Such documents are never seen in their complete form on most computer systems. In addition, documents were also extended with commands to execute programs needed by the browser. However, this feature quickly ran into trouble. The code plugged into the document by the author was specific to the local processor and operating system and therefore could not be executed in the browser's system. This was perhaps because of a different processor architecture or operating system version or the runtime libraries. In addition, there was the hazard of potentially executing malafide pieces of code called computer viruses embedded within the documents, causing problems to the recipient system and spreading it around. Thus the Internet has thrown up a number of technological challenges like using the net for secure, private and fast commerce as well as supporting video and 3D graphics. Yes, the web has now become the all-pervasive computer! And the *web page* is no more a passive entity.

Given this scenario of the Internet, there was an implicit need to come up with a programming language to enable authors to incorporate program application fragments or *applets* into the web document. It had to be sophisticated, system independent and secure so that the author concentrates on the application and is not concerned with its execution.

The Programming Language Java

The *Java* programming language was designed to cater to the above mentioned requirements, overcome a number of problems in modern programming practice and enable the fast development of distributed programs and applets. It is a simple, object-oriented language that has been designed to be portable, robust, secure and efficient. It's compiler, runtime system and class libraries have been targetted for most processor architectures and

What is Hypertext?

All books and articles are linear documents where the order of information presentation is fixed by the author or editor. This article and the journal are typical examples. Generally the title page is followed by a contents page and then by the articles. Finally we have the references and the index. While browsing through a book we cannot obtain information in any other order of presentation. However, the computer permits one to create documents in the system wherein multiple presentation orders can be designed. The person browsing through it can choose a suitable order. Thus the organization of information becomes non-linear and multi-dimensional. Such texts have *hyperlinks* or cross-references between different parts. Hence they are called *hypertexts*.

Browsing through a Hypertext Document on the Web

Assume that you have a magic wand and table and are interested in learning how a computer works. You ask for a computer book and wave your wand; and lo and behold, there is the book on computers on the table opened at the page which explains how it works. You read a paragraph in between and see a reference to a computer journal paper. You want it and with a wave of your wand, the referenced document, opened at the topic of interest appears stacked over the book on your table. You start reading this paper and come across another reference, now to an encyclopedia. You wish for it and with a wave of your wand, the third document appears, stacked over the other two and opened at the topic of your interest. You start reading it. You come across a reference to a figure explaining the electronic circuits comprising the CPU. You ask for it and you get a sheet with the figure in all colours and precision. You finish noting the intricacies of the circuit diagram and wish to get back to the earlier document. With a wave of your wand, the detailed figure disappears and you are left with the encyclopedia you were reading. You find it becoming irrelevant, so with a wave of your wand you get back to the journal paper. You begin to get more curious and seek clarifications. You wish you could ask the authors; you draft a letter and with a wave of your wand, send it off. Not wanting to get into further details of the paper, you wish to get back to the initial computer book you started with. Before doing that you wish to make a note (or put a bookmark) at the journal paper. Again with a wave of the wand, you get it conveniently recorded. Finally, you observe that the journal paper disappears and you are left at the right place in the first book that you started reading on how the computer works. A *web* user has a similar experience where the magic wand is replaced by the mouse and the magic table is your computer monitor.

operating systems. The class libraries are rich and powerful, enabling the quick and effective development of distributed web applications.

Java is a simple, object-oriented language that has been designed to be portable, robust, secure and efficient.

Java is a derivative of C / C++: simple, familiar syntax and with fewer complex features. It does not have many of the poorly understood, confusing and rarely used features of C++. There are no pointers in Java — only tightly bounded arrays. This eliminates the possibility of overwriting memory and corrupting data unwittingly. Pointers have been the source of raw power in C and C++ programs as well as the primary feature that helped introduce bugs into almost all programs. Java is sophisticated enough to help programmers express complex ideas easily, using

arrays and has an efficient *automatic garbage collector* based memory management scheme. In addition, unlike C or C++, Java does not support *structures, enums* or *functions*. Every programmable object is an instance of a 'class' in object-oriented Java. This class definition permits both *static* and *dynamic inheritance* and therefore full reuse of code. A *class* is a template that characterizes an object's internal representation as well as its behaviour. The internal representation subsumes the role of a structure in C or C++. The behaviour of an object is essentially captured in terms of *member-functions* or *methods* in the class definition. In addition, an object template, that is a class, can also be defined in terms of other class definitions or properties or behaviour that can be inherited dynamically (or statically) if needed. Java is truly object-oriented in that many class definitions can be inherited dynamically. This combined with the inheritance of programs across the network from remote systems enables applets to be inherited during runtime. However Java does not support multiple inheritance whose semantics and usage has been quite controversial in other languages. Following the principles of structured programming, there are no *go to* statements, no automatic type-casting or operator overloading. Java is so versatile that one of the big programs first written using it, was its own interpreter and compiler.

Why is Java called Java ?

Java designer Jim Gosling couldn't decide on a better name than that of an oak tree growing outside his office window and called it 'Oak'. However it ran into trademark problems for Sun Microsystems. After a bit of a search he called it Java, slang for the best coffee that was imported into California from Indonesia decades ago. Initially Oak was aimed at programming a heterogeneous network of electronic home appliances. This defined Java's characteristics as a small, reliable, real-time system.

The most interesting contribution of Java has been in its runtime system: support for automatic memory management and support for *multi-threading*. The Java garbage collector keeps track of all objects generated, automatically freeing the memory used by objects that have no further use and are not referred to by other existing objects. This contributes enormously to making the code robust. In addition, the support for multi-threading enables the efficient execution of programs that potentially have multiple threads of control. Thus the garbage-collector is efficiently run as a background process overcoming the biggest drawback of garbage collector based languages. Multi-threading is supported via inheritable thread class libraries. It results in better interactive response and near real-time features. This enables Java to inter-

Java programs are portable because they can be executed on any system without change.

face and support many features of modern operating systems and network protocols.

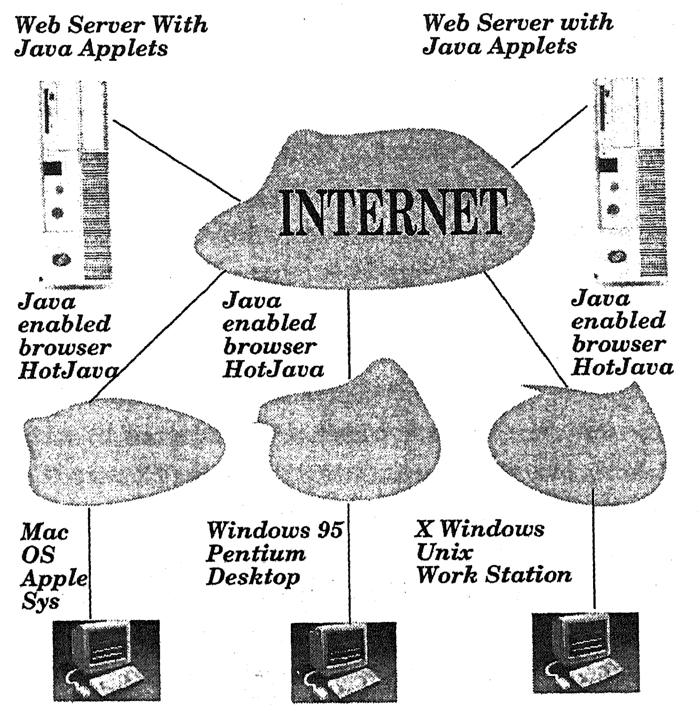
Java programs are portable because they can be executed on any system without change. The Java compiler transforms the program sources into the instructions of an abstract processor called the *Java virtual machine*. Thus the compiled Java code is architecture neutral. The *bytecode instruction interpreter* specific to each processor type first verifies the consistency of the compiled program and then efficiently converts this code into the native processor instructions before executing it. Java programs are robust because explicit memory manipulations by the programmer are prevented: memory addresses cannot be *dereferenced* nor can *pointer arithmetic* be used to access object contents. Array bounds are checked so that array indices are never out-of-bounds. Java programs are secure: distributed applications have to exhibit the highest levels of security concerns. A *bytecode verifier* in the Java interpreter ensures that the compiled code is strictly language compliant, thus trapping all malafide modifications, more so the computer viruses parading as legal code. In addition, by making the Java interpreter determine the memory layout of all objects at runtime, all possible means of infering memory contents at compile time and appropriately accessing them at runtime are eliminated — these are the potent security holes in most executables. Using conventional languages to come up with such distributed applications as rigorous as Java has been difficult. It is no wonder that Java is popular with Internet applet programmers.

What is a URL ?

A typical website address for a hypertext document consists of four parts. For example, this article is available over the internet at: http://serc.iisc.ernet.in/~mohan/java-art.html The first part ('http:') is the protocol to be followed in accessing the document. The second part ('serc.iisc.ernet.in') gives the domain address of the system which has this document. The third part ('~mohan') indicates that the document is in the home directory of the user with a login 'mohan', while the last part gives the filename of the document ('java-art.html'). Such addresses are also called uniform resource locators (URLs) in the web terminology. Java applications can open and access documents and objects across the network using these URLs with the same ease as accessing a local file.

Java And Internet

The Java programming language enabled the web document authors to deliver small application programs to anyone browsing the pages of the html documents. The page became alive because it could create game scoreboards, execute animated cartoons, audio files and video clippings. In addition, it changed the way Internet and world wide web worked by allowing archi-



Java enabled browser gives a uniform and universal interface to the World Wide Web over the Internet.

tecturally neutral compiled code to be dynamically loaded from anywhere in the network of heterogenous systems and executed transparently.

One of the popular web browsers is the *HotJava* entirely written in Java. This browser brings out the best of Java and Internet programming. It incorporates the Java runtime system and thus enables one to execute Java applets embedded within html documents. The potential applications which can be written using Java are enormous: Secure commercial transactions spanning multiple nations and continents are easy, more so those using electronic cash. Based on it we can develop applications whose usage can be metered across the network — the latest version being available from the nearest easily accessible host

Making an Active Web Document

An active Java or html document is essentially a text file with html commands for accessing links, displaying various text strings in different fonts and colours as well as graphical objects. This is along with support for interactive execution of a variety of application program fragments. This makes the document come alive with video and sound while browsing. Thus a book on Assamese folk music can have music programs or applets embedded in the document. When you wish to hear a representative piece, you can execute and hear it on your system. Writing such an applet involves digitizing the musical piece, storing it compactly while retaining the original fidelity, using appropriate class libraries to interface with the sound subsystem hardware and interfacing with the browser's interactive command interface.

(which may change with every access depending on the network traffic conditions). *Intelligent Agent* applets can be programmed so that they can comb the network for the latest and most useful information as well as in doing routine decision making (like filtering out junk information) and chore handling (like routine backups or monitoring particular sites for information alerts). The size of the basic interpreter is about 40KB and the standard libraries for multi-threading add another 170KB.

In the near future, a sizeable part of our professional activity will be over the Internet, and our success depends on our ability to adapt to and use the Internet technologies.

While the clout and support of Sun Microsystems has been the main reason for the success and popularity of Java and the Hot Java browser, the ease with which it was adopted by various programmers over the Internet forced rivals like Netscape Communications and Microsoft to accept it. The latest version of the *Netscape Navigator* web browser supports Java applets. In addition, Netscape Communications came up with an augmented design of Java called the *Javascript* language. However, Microsoft has been pursuing a project code named *Blackbird* that is expected to be a strong competitor to Java and its derivatives. Blackbird is a complete authoring and publishing environment that harnesses all the features of *Windows 95* and *OLE2.0*. Internet technology watchers predict a product war in the near future wherein Blackbird applications over the Internet, more so in the world of business and financial transactions, are expected to dominate and perhaps almost kill competitive products, something like what Windows 95 and Microsoft Office have done to their contemporaries today. But one outcome is certain — the Web has become the computer — and costly desktop systems will soon be replaced by a simple computer connected to the Internet, costing less than Rs 20,000 and capable of executing any program dynamically loaded from anywhere in the world. In this case, companies will sell computing capabilities (viz., number crunching, visualization, data mines, information warehouses) over the Internet cable even as our electricity boards supply us with power, albeit more reliably. This means that in the near future, a sizeable part of our professional activity will be over the Internet, and our success depends on our ability to adapt to and use the

Internet technologies. If the necessary infrastructure gets in place soon, India will have caught up with the rest of the world and will have arrived in the *global information society*.

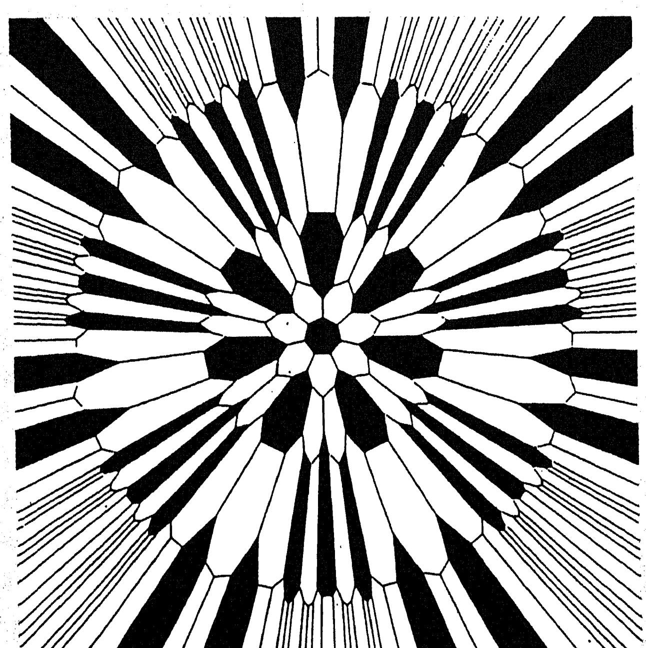
How to Learn More about Java

The best source for learning more about Java is the Internet which has many Java application servers: Access the website: http://java.sun.com and you can get all documents and the software (including a HotJava browser, a Java compiler, class libraries etc) for free. In addition you can join electronic mailing lists and newsgroups to interact with others around the world who work on Java. The book Java! by Tim Ritchey, New Riders Publishing, Indiana, 1995 is one of the first few books on the language.

Suggested Reading

Tim Ritchey. Java!. New Riders Publishing, Indiana, USA. 1995.

Address for correspondence
T S Mohan
Supercomputer Education and Research Centre
Indian Institute of Science
Bangalore 560 012, India.



Tesselation of convex heptagons

Molecule of the Month

Maitotoxin - Holder of Two World Records

J Chandrasekhar

J Chandrasekhar is at the Department of Organic Chemistry of the Indian Institute of Science, Bangalore.

The structure of a large and extremely toxic natural product is described.

Nature is an excellent synthetic organic chemist. Using mild reaction conditions and a few elemental combinations, a large variety of complex molecules are made in and around us. The atoms are put together in precise arrangements to enable the molecules to carry out different tasks with remarkable specificity. The treasure house of natural chemicals contains delicate perfumes, spectacularly coloured substances, medicines for numerous ailments and also the deadliest of poisons. It is a stimulating and challenging exercise to determine the molecular structures of various substances and to relate them to their properties. From a chemical point of view, even poisonous compounds are interesting.

The most venomous substances associated with snakes, scorpions, and even some bacteria and plants have a common structural feature. These compounds are in fact proteins. However, many non-peptide toxins are also known. Palytoxin was considered the most toxic of such compounds till recently. It is one of several poisonous marine natural products. A compound named tetrodotoxin, found in puffer fish, is another well known example. It was shown to be the culprit in numerous cases of food poisoning associated with consumption of sea food. The molecule which should perhaps head the list of deadly natural toxins is maitotoxin, named after the Tahitian fish maito.

From a chemical point of view, even poisonous compounds are interesting.

Maitotoxin was first discovered in 1976 from the surgeon fish *Ctenochaetus striatus*. It has also been isolated from cultured cells

of *Gambierdiscus toxicus* and has been purified using HPLC.[1] Toxicological investigations and physicochemical studies to determine the molecular structure have been carried out on purified samples. Maitotoxin turns out to be remarkable both for its structure and its biological activity.

[1] High performance liquid chromatography is an excellent procedure to separate components from a mixture. It is an essential tool in natural products research.

Maitotoxin is lethal. In the units normally used to specify toxicity, the LD_{50} value (lethal dose to kill 50% of subjects) is 50 nanogram (10^{-9}g) per kilogram of mice. Since a mouse usually weighs no more than 20 g, one can state the result of the experiment in a macabre fashion: a gram of maitotoxin can kill approximately half a billion mice! No non-peptide natural product has such lethal potency.

Even to the chemists who abhor violence, maitotoxin is appealing for a different reason. It is the largest molecule made by nature, leaving aside bio-polymers like polypeptides and polysaccharides. The molecular weight of maitotoxin is 3422 Daltons.[2] Thus, the molecule currently holds two world records: it is the largest non-biopolymeric natural product and the most lethal non-peptide natural product.

[2] One atomic unit of mass is called a Dalton.

The full 3-dimensional structure of maitotoxin has not yet been solved using X-ray crystallography. But the basic molecular skeleton as well as the complete stereochemistry (relative disposition of atoms at chiral centres) have been fully worked out. Using infrared, ultraviolet and mass spectra, many key features of the structure were determined. More details of the structure could be obtained only through nuclear magnetic resonance spectroscopy, especially through extensive two-dimensional (2D) NMR techniques. Molecular mechanics calculations were also imaginatively used to interpret the coupling constant patterns as well as Nuclear Overhauser Effect data.[3] It is truly a triumph of modern spectroscopic methods that the stereo-structure of a molecule with such complexity can be solved.

[3] Vicinal coupling constants are related to H-C-C-H dihedral angles through a simple expression called Karplus equation. NOE data provide information about distances between non-bonded hydrogen atoms.

The gross structure of maitotoxin is shown in *Figure* 1. The

molecule contains numerous fused saturated rings. To keep track of different portions of structures of polycyclic molecules, organic chemists usually label the rings A, B, C, etc. The rings in maitotoxin exhaust the English alphabets and so the labelling has to go beyond Z to A', B' and so on up to F'. Note the presence of numerous oxygen atoms. There are 32 ether linkages and 28 hydroxyl groups. These are evidently involved in the interactions of the molecule with cations. It is not a coincidence that the biological activity of maitotoxin is linked to its ability to elevate intracellular Ca^{2+} concentration. The molecule is likely to be an excellent model for probing cellular events associated with Ca^{2+} flux.

Maitotoxin currently holds two world records: it is the largest non-biopolymeric natural product and the most lethal non-peptide natural product.

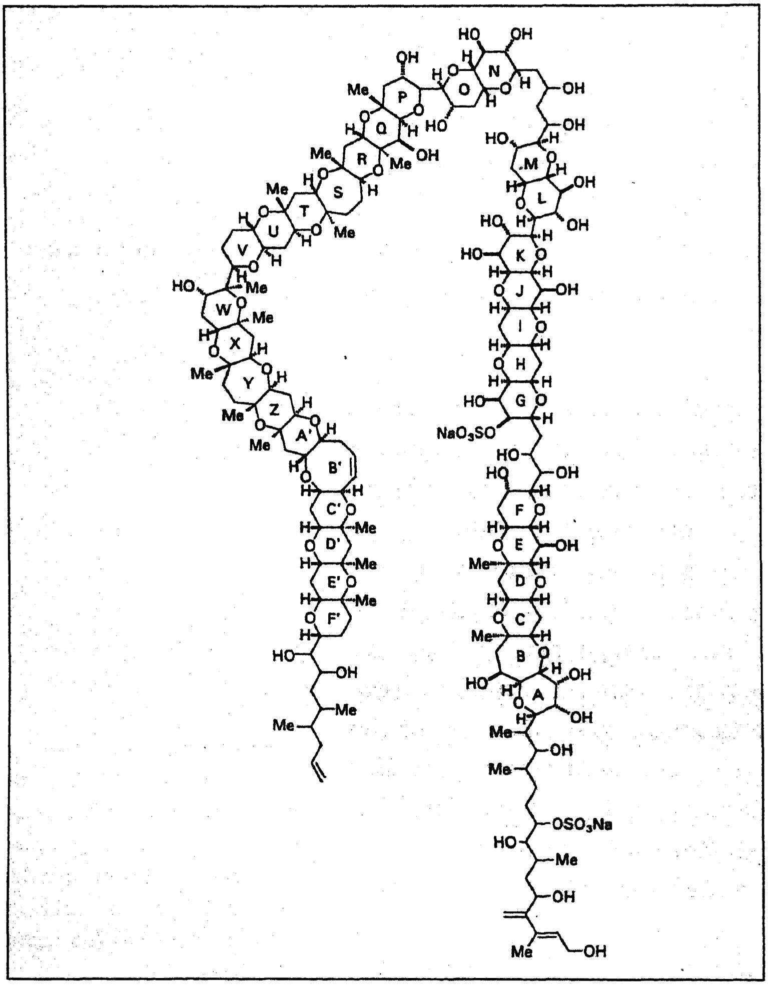
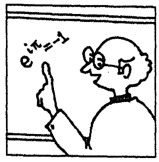


Figure 1. Molecular structure of maitotoxin. Not all stereochemical relationships are shown.

Address for correspondence
J Chandrasekhar
Department of
Organic Chemistry,
Indian Institute of Science,
Bangalore 560 012, India.



In this section of **Resonance**, *we invite readers to pose questions likely to be raised in a classroom situation. We may suggest strategies for dealing with them, or invite responses, or both. "Classroom" is equally a forum for raising broader issues and sharing personal experiences and viewpoints on matters related to teaching and learning science.*

! Pressure Melting and Ice Skating

In cold countries, ice skating is a popular sport. The skater is supported on two metal 'blades' which move over the ice with amazingly low friction, allowing very rapid and graceful movements (and disastrous falls!). For generations, the text book explanation of this low friction has been the fact that ice melts under pressure (most other substances freeze under pressure). The extract given below, from a recent article by S C Colbeck in *American Journal of Physics* (63 (10): 888 October 1995) tells us that the physics of ice skating is richer, and hence more interesting, than mere pressure induced melting.

> "While pressure melting is commonly thought to be the mechanism responsible for the low friction of ice, there are many arguments against it. The high pressures required would cause failure of the ice unless it is well confined by the blade. If the ice and meltwater are confined, just below -20° C liquid water cannot coexist with ice at any pressure because the high-pressure forms of ice appear. If the mechanism does operate, the high pressures would cause high rates of water loss by squeeze in very thin films. Frictional heating by shearing of the water films is much more plausible, as shown quantitatively by the example with equal contributions by the two mechanisms".

The physics of ice skating is richer, and more interesting, than mere pressure induced melting.

Ravi Divakaran, Department of Chemistry, St. Albert's College, Eranakulam 682018, Kerala.

? Bunsen Burner- Revisited

This is in response to the article titled "On Bunsen Burners, Bacteria and the Bible" written by Milind Watve which appeared in the "Classroom" section of ***Resonance***, February 1996. As a teacher of chemistry, I read the author's observations on the textbook experiment to determine the percentage of oxygen in air, with great interest. I carried out this experiment carefully, as follows:

What immediately occurred to me was that drinking glasses usually have a tapered shape and therefore equal distances marked off on its wall may not represent equal volumes. Secondly, the candle inside the glass will take away some volume and therefore I assumed that introduction of a second candle, as reported by Watve, must have decreased the effective volume further and contributed to the increase in water level noted.

Luckily, I could get hold of a fancy, glass tumbler with vertical walls, which I used for my experiments. In order to eliminate the differences in volume due to the candlesticks, the following procedure was adopted. Three candle sticks with heights about half that of the tumbler were taken. These were placed inside the tumbler which was then filled with water. The candles were then taken out and the level of water inside the tumbler fell. The volume of water now represents the effective volume inside the tumbler when the three candles are covered. This height was marked off from the open end of the glass and divided into 10 equal parts to provide 10%, 20% . . . 100% readings by sticking a strip of paper along the wall with zero at the open end.

It appears that the decrease in the volume of air inside the glass depends on the number of candles lit (about 15% for each candle lit) and not on the oxygen content.

The three candles were stuck to the bottom of a trough and all were always present during the experiments so as to avoid any volume changes due to the number of candles present. One glass of water was then poured into the trough. One candle was lit and covered with the glass tumbler carefully. When the candle was extinguished, water rose in the tumbler and occupied about 15%

of the effective volume after a few minutes when the apparatus cooled down to almost room temperature. This was repeated several times, each time replacing the water at the bottom of the trough with a similar fresh quantity in order to avoid any effects of dissolved CO_2 etc., and lighting only one candle. Each time, the water level inside rose to occupy about 15 to 18% of the effective volume.

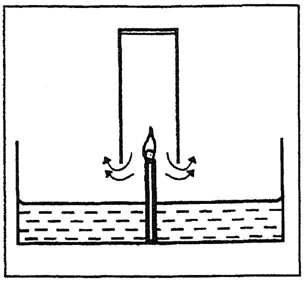


***Figure 1** Air escapes at the bottom when a tumbler is lowered over a candle in a trough of water.*

The experiments were then repeated, this time lighting two candles at a time. Water level in this case rose to occupy nearly 30% of the effective volume. The experiments were then repeated by lighting all the three candles. It was observed that the water level now reached about 45% of the effective volume. It thus appears that the decrease in the volume of air inside the glass depends on the number of candles lit (about 15% for each candle lit) and not on the oxygen content.

One possible explanation for the phenomenon may be as follows. One invariably takes a few seconds to lower the tumbler over the candle, during which time the air inside it expands and escapes at the bottom as indicated in *Figure 1*. When the mouth of the glass is closed by the water level and the candle stops burning, the air inside cools and contracts. Thus the rise in water level inside the glass represents the air that escaped due to expansion. When more candles are lit, the heat increases proportionately, driving away a proportionate amount of air from the tumbler. It so happens that the expansion loss per candle flame is about 15 to 18% and is close to the estimated oxygen content of air when only one candle is used. Water level should rise slowly during the experiment while the candle is still burning if it is due to the oxygen being used up. It was observed that there was little rise while the candle was burning, but the water level rose all of a sudden as soon as it was extinguished.

It so happens that the expansion loss per candle flame is about 15 to 18% and is close to the estimated oxygen content of air when only one candle is used.

Further, when hydrocarbon in the candle burns, $C + O_2 \rightarrow CO_2$ (neglecting the volume of water produced as it is liquid) for every molecule of oxygen used, one molecule of CO_2 is produced and

When hydrocarbon in the candle burns, for every molecule of oxygen used up one molecule of CO_2 is produced and therefore there can be no change in volume.

there can be no change in volume. Indeed this has been observed to be the case when charcoal is burned inside a gas jar inverted over mercury. When water is used, it may absorb a portion of the CO_2 because of its greater solubility compared to air. It has also been documented that when a burning splinter is extinguished inside a closed gas jar, the remaining air contains only 2.5% CO_2 and 17.5% oxygen still remains!

Thus it appears that the experiment which has been taught to students over all these years does not represent the oxygen content of air at all!

Suggested Reading

J R Partington. A Text-Book of Inorganic Chemistry. 6 ed. ELBS. 1963 Especially see pages 620 and 622.

G S Ranganath, Raman Research Institute, Bangalore

? In introducing the Bohr theory of the hydrogen atom, one makes a postulate that electrons in certain special orbits around the centre do not radiate. How does one reconcile this with what students have already learnt about radiation?

Discussion of questions raised in the Classroom section of **Resonance** *Vol. 1, No.1.*

From Bohr's theory we can calculate $\nu = (E_n - E_{n-1})/h$ the frequency of electromagnetic radiation emitted by the atom in a transition from n^{th} to $(n\text{-}1)^{th}$ state. For large values of n we find (verify) the radii and the energies of the n^{th} and $(n\text{-}1)^{th}$ states to be very nearly the same and ν becomes equal to the frequency of the orbital rotation of the electrons. This is the result to be expected from electrodynamics. Hence at large quantum numbers Bohr's theory agrees with classical mechanics. Incidentally, this is the central message of Bohr's Correspondence Principle.

Mohan Devadas, SBM Jain College, Bangalore.

? As an example of special relativity in action, one quotes the case of the muon, with a half-life of less than ten nanoseconds (1 nanosecond = 10^{-9} sec). Travelling at almost the speed of light, it should only be able to cover a few metres in this time.

But cosmic ray physicists are able to detect muons which have travelled several kilometres, from the top of the atmosphere. Is this an example of length contraction or time dilation?

For an observer on earth it is an example for time dilation. However, for an observer on the muon it is an example for length contraction. Hence the answer is observer dependent.

? Carbon monoxide has a small dipole moment with the negative end at carbon. How can one explain this result?

J Chandrasekhar, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012.

In a heteronuclear diatomic molecule, A-B, the charge distribution would be unsymmetrical. The bond(s) would be polarised such that there is greater electron density near the atom with the higher electronegativity, say B. Atom B would therefore have an excess negative charge (δ–) with a corresponding positive charge on A. The dipole moment of the molecule would then be Rδ, where R is the internuclear separation or bond length in A–B. The negative end of the dipole is obviously B.

Discussion of questions raised in the Classroom section of **Resonance** *Vol. 1, No.2.*

Using the above arguments, the dipole moment of CO is expected to be fairly large with the negative end at oxygen. The experimental finding is different. One possible interpretation is that the canonical form (or resonance structure) with a negative charge on carbon makes a significant contribution to the electronic structure of CO.

$$C = O \;\longleftrightarrow\; {}^{-}C \equiv O^{+}$$

The above explanation is only partly correct. There is a more important reason for the failure of the qualitative arguments. An assumption made in the preceding description of dipole moments is incorrect. The excess electron density was assumed to be a point charge at the nucleus. This is reasonable only if the charge distribution is symmetrical about the atoms. In reality, the electron density distribution is considerably uneven. This has to be taken into account in order to get the correct estimate of the dipole moment of a molecule.

The centroids of electron densities in 'pure' atomic orbitals such as s, p, d ... are at the corresponding nuclear centres. However, this is not the case in hybrid orbitals. One can readily see from the shapes of hybrid orbitals (e.g. *sp* orbital) with one lobe larger than the other that the centroids are shifted away from the nucleus. While computing the electronic contribution to the dipole moment, the magnitude of the charge as well as the average location have to be considered. The correction for the shift in the electronic centroid is sometimes called the hybridisation contribution to the dipole moment. The total dipole moment is the net sum resulting from positive nuclear charges, contribution from effective negative charges assuming the electron densities to be centred around the nuclei and a hybridisation correction. The individual terms cannot be experimentally measured. But they can be calculated using quantum chemical methods like molecular orbital theory. The computed data give us insights into the nature of bonding in molecules.

The total dipole moment is the net sum resulting from positive nuclear charges, contribution from effective negative charges assuming the electron densities to be centred around the nuclei and a hybridisation correction.

From such calculations, we get the following bonding description. The σ and the two π bonds of CO are polarised more towards oxygen, as expected. There are two other filled valence orbitals. For the sake of simplicity, they may be viewed as sp hybridised lone pair orbitals on oxygen and carbon. The centroids of these orbitals are located beyond the CO molecule, significantly away from the corresponding atoms. The centroid of the carbon lone pair is displaced to a greater degree. The resulting hybridisation correction is so large in this molecule that it reverses the trend from the effective charges on the atoms. The total value of the dipole moment is small and the carbon atom is at the negative end.

The consistency of the above description can be verified by considering the dipole moments of other molecules containing the C=O fragment. In molecules like formaldehyde or acetone, there is no lone pair on carbon. The total dipole moment is determined primarily by the effective charges on the atoms. The

xygen atom is the negative end of the dipole in these type of ıolecules, as expected from electronegativity of the atoms ıvolved.

Rajaram Nityananda, Raman Research Institute, Bangalore 560 080.

In a laser beam, many photons occur with the same irection, frequency and polarisation. Is this an example of ose-Einstein condensation (BEC)?

. laser beam indeed has many photons in the same quantum :ate, and these obey Bose statistics. However, Bose-Einstein ɔndensation is an equilibrium phenomenon in which the particles re given enough time to exchange energy with each other or with ɔme other system. The situation in a laser is quite different. The ctive medium is pumped, i.e driven to excited energy state, by source of optical, electrical, or even mechanical energy. It then roduces the laser beam which escapes, surely a highly non-quilibrium situation.

[owever, if we allow photons to come to equilibrium with each ther, the result is black body radiation. The distribution of articles as a function of energy is smooth, and does not have the pike at zero energy which is the signature of Bose-Einstein ɔndensation. The reason is that as we cool the box the photons an disappear by absorption in the walls. The mean spacing etween particles increases. This offsets the effect of the increase ı the de Broglie wavelength on cooling. In the case of an rdinary gas the number of particles is fixed and hence also the ıean spacing. (see the Research news item in the February 1996 ;sue). As we cool, we get condensation.

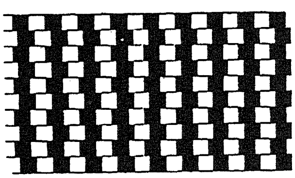
Rajaram Nityananda, Raman Research Institute, Bangalore 560 080.

A collection of non-interacting Bose particles exhibit BEC t low temperatures. How is this possible in the absence of ıterparticle forces? One usually assumes that an ideal gas oes not condense.

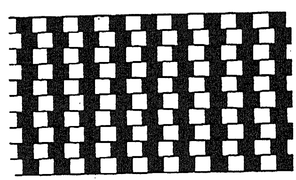
'he question of how particles with no forces between them can

condense is indeed a deep one. The rules of quantum statistics tell us that the presence of one particle in a state can forbid (Fermi) or encourage (Bose) the presence of a second identical particle, which therefore "knows" about the existence of the first. While these rules of quantum theory have been known for a long time, there is no deeper model of this behaviour. The fact that a quantum system must be regarded as a whole, even when the parts are non-interacting, is called "non-locality" and it appears in other situations as well, eg. the famous EPR (Einstein, Poddsky, Rosen) paradox.

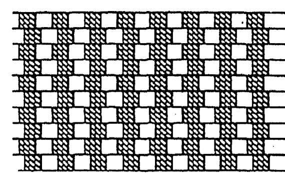
Optical illusions ...



(1)



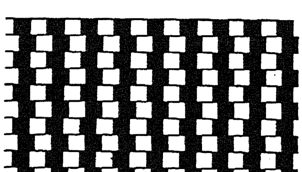
(2)



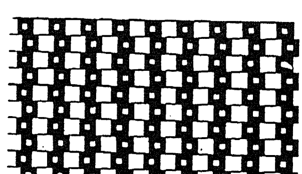
(3)

We see that the rows seem to alternatively converge to the right and left edges. This illusion is sensitive to the contrast and the pattern of the repeat unit. It is absent when there are grey squares as seen in No.3 above. When we compare Nos. 4, 5 and 6 the illusion appears to be greatly diminished in No.6 where the white dot is off centre. (T N Ruckmongathan, Raman Research Institute.)

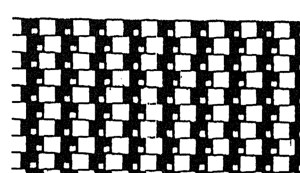
(4)



(5)



(6)



Think It Over

This section of Resonance *is meant to raise thought-provoking, interesting, or just plain brain-teasing questions every month, and discuss answers a few months later. Readers are welcome to send in suggestions for such questions, solutions to questions already posed, comments on the solutions discussed in the journal, etc. to* Resonance, *Indian Academy of Sciences, Bangalore 560 080, with "Think It Over" written on the cover or card to help us sort the correspondence. Due to limitations of space, it may not be possible to use all the material received. However, the coordinators of this section (currently A Sitaram and R Nityananda) will try and select items which best illustrate various ideas and concepts, for inclusion in this section.*

1 Capillarity

From Rajaram Nityananda, Raman Research Institute, Bangalore.

We are all taught the phenomenon of capillarity, in which a liquid rises in a narrow tube, when it 'wets' the material of the tube. The formula for the height is $h=2T/(r\,\rho\, g)$. T is the surface tension of the liquid, r the radius of the capillary tube, g the acceleration due to gravity, and ρ the density of the liquid. Here are two questions relating to this everyday phenomenon.

1) What happens if the height of the tube is less than the value given by the formula above? Would the liquid squirt out? (Beware of perpetual motion!)

2) The formula contains the quantity T, which is a property of the liquid, but does not appear to contain any property of the material of which the capillary tube is made. This is surprising, because surely it is the attractive force between the material of the tube and the liquid which is responsible for the liquid rising against gravity! What is going on?

From Rajeeva L Karandikar, Indian Statistical Institute, New Delhi.

2 To Switch or Not to Switch

You are a winner in the preliminary round of a TV game show and the host gives you a chance to win the super prize: a fancy car. You are shown three doors numbered 1, 2 and 3. Behind one of them is the car. You are asked to choose a door. If the chosen door is the one hiding the car, you win the prize.

You choose, say, door number 2. The host of the show then says: "First, let us see what is behind door number 1?" He opens it and you see that the car is not there. Now he asks you: " Do you want to stay with your initial choice (number 2), or would you like to switch to door number 3?" What would you do?

Does this have a familiar ring to it? May be the 'Prisoner's dilemma' has the same logic.

Mohan Delampady, Indian Statistical Institute, Bangalore.

3 Prisoner's Dilemma

Discussion of question raised in **Resonance** *Vol. 1, No.2.*

Three prisoners, A, B and C are each held in solitary confinement. A knows that two of them will be hanged, but one will go free. However, he does not know who will go free. He thus reasons that there is a 1/3 chance of his survival. Anxious to know his fate, he asks his guard. But the guard will not tell A his fate. A thinks and puts the following proposal to the guard: "If two of us must die, then I know that either B or C must die and possibly both. If you tell me the name of just one of them who is certain to die, then I learn nothing about my fate; and since we are kept apart, I cannot inform them of theirs. So tell me which one of B or C is to die?" The guard accepts the logic and tells A that C is to die. A now reasons that either he or B will live. Thus A now has a 1/2 chance of survival. Is A's reasoning correct?

A's argument is incorrect. He hasn't considered what the guard might say if both B and C are to die. Conditional probability argument is needed to analyze this puzzle. Let AB be the event that A and B are to die and define BC and AC similarly. Let D be the event that the guard says 'C is to die'. Then

$$P\,(A\text{lives}\,|\,D) = P\,(BC\,|\,D) = \frac{P\,(BC\text{ and }D)}{P\,(D)} = \frac{P\,(D\,|\,BC\,)P(BC\,)}{P\,(D)}$$

Note that

$$P(D)=P(D|AB)P(AB)+P(D|BC)P(BC) + P(D|AC)P(AC)$$
$$=P(D|BC)P(BC) + P(AC)$$

since $P(D|AB) = 0$ and $P(D|AC) = 1$. Therefore,

$$P(A\text{lives}|D) = \frac{P(D|BC)P(BC)}{P(D|BC)P(BC) + P(AC)}.$$

Since A initially believes that they all have an equal chance of surviving, it follows that $P(AB) = P(AC) = P(BC) = 1/3$. If he further assumes that the guard is equally likely to say 'B is to die' and 'C is to die' if BC is to occur, then $P(D|BC) = 1/2$. Therefore,

$$P(BC|D) = \frac{(1/2) \times (1/3)}{(1/2) \times (1/3) + (1/3)} = \frac{1}{3}.$$

That is, his probability of survival should be 1/3 still. Note, however, that he might not hold $P(D|BC) = 1/2$. Then other values of $P(BC|D)$ are possible. When does he obtain $P(BC|D) = 1/2$?

Count Rumford burns his tongue ... Water was considered a good conductor of heat until, in the last years of the eighteenth century, Count Rumford burnt his mouth when eating apple pie; he then decided to do some experiments 'to show that water, and probably all other liquids, are non-conductors of heat'. (From *Stories from Science IV* by Sutcliffe and Sutcliffe)

Evidence for Bird Mafia!*

Threat Pays

Raghavendra Gadagkar and Milind Kolatkar

Birds are remarkable for their extraordinary efforts at nest building and brood care. Given that so many species of birds spend so much time and effort at these activities, there is plenty of room for some species to take it easy, lay their eggs in the nests of other species and hitch-hike on their hosts. The cuckoo that lays its eggs in the nests of a variety of host species is well known. Indeed, over 80 species, i.e., over 1% of bird species are known to be such obligate inter-specific *brood parasites.* These include two sub-families of cuckoos, two types of finches, the honey guides, the cowbirds and the black-headed duck. Because parasite species often use more than one host species, more than 1% of bird species act as hosts to brood parasites. Inter-specific brood parasitism has evolved independently at least seven times in birds and can have a significant effect on the populations of the host species and even lead to their extinction. Although hosts sometimes detect and eject alien eggs, their success in ridding their nests of parasite eggs is often very limited and that is why brood parasitism has survived as a way of life. One reason for such limited success of the hosts is the exquisite mimicry often exhibited by the parasites whose eggs are virtually indistinguishable from those of the host. What is perplexing however is that many parasite species lay eggs that look nothing like their host's eggs and yet get away with it. Obviously hosts have not perfected the art of removing all or most of the alien eggs. But why should this be so?

* Abridged version of an article which appeared in *Current Science* — reprinted with permission.

Over 1% of bird species are known to be obligate inter-specific *brood parasites.*

Amotz Zahavi has suggested the hypothesis that parasites such as cuckoos may repeatedly visit the parasitized nests and destroy the eggs of the host if it has ejected the parasite's eggs and not do so if the host has accepted them and is taking good care of the parasite's eggs/chicks. In the presence of such a parasite 'Mafia', hosts who are incapable of defending themselves against the attacks of the parasites may find it better to accept some parasite eggs and additionally rear at least some of their own rather than lose all their eggs in the parasite attack. There has recently been an attempt to test this Mafia hypothesis using the great spotted cuckoo *Clamator glandarius* and its host the black-billed magpie *Pica pica* in Spain. There is evidence that cuckoos visit nests where they have laid eggs and peck at magpie eggs if their own are missing. A magpie

A magpie has three options - rear both magpie and cuckoo chicks, eject the cuckoo eggs and rear only its own offspring or abandon the nest altogether and start all over again.

that finds cuckoo eggs in its nest appears to have three options - accept the parasite's eggs and rear both magpie and cuckoo chicks, eject the cuckoo eggs and rear only its own offspring or abandon the nest altogether and start all over again.

When magpies ejected cuckoo eggs, 86% of their nests were attacked by the cuckoos but when they accepted cuckoo eggs, only 12% of their nests were attacked, a difference that is statistically significant. Predation rates were of the order of 22% in non-parasitized nests. All magpies re-nesting after loss of eggs to cuckoo predation accepted cuckoo eggs without ejecting them or abandoning their nests in the second breeding attempt. But did the cuckoos destroy magpie eggs just to get the magpies to re-nest and provide another opportunity for them to lay their own eggs? If inducing the magpies to re-lay was the main objective, magpie nests, early in the season (which have a higher probability of re-nesting) rather than those late in the season (which have a substantially lower probability of re-nesting), should suffer higher rates of attack by the cuckoos. However, late nests suffered a slightly higher rate of predation compared to early nests.

Magpies that accepted the cuckoo eggs produced 0.43 ± 0.10 (mean s.d.) fledglings per nest, while those that ejected the cuckoo eggs produced 0.29 ± 0.29 fledglings per nest and finally, those that abandoned their nest and started all over again produced 0.40 ± 0.31 fledglings per nest. The measured reproductive success of the abandoners should be halved at least, because the probability of recruitment of offspring into the breeding population decreases dramatically as the season progresses, thus giving us a figure of about 0.20 fledglings per nest for the abandoners. In other words acceptors, ejectors and abandoners have about the same reproductive success values that are not significantly different statistically.

A more powerful approach is to experimentally remove cuckoo eggs from some parasitized magpie nests and do no such thing in a group of control, parasitized nests. When this was done, nests from which cuckoo eggs were experimentally removed (equivalent to ejectors) produced 0.85 ± 0.28 fledglings, while the control nests (equivalent to acceptors) produced 0.54 ± 0.24 fledglings per nest. These numbers are also not significantly different statistically. Does not the lack of significant differences between the acceptors, abandoners and ejectors in the natural population and the experimental and control nests in manipulated samples weaken the Mafia hypothesis? Not really; it would be naive to expect the Mafia to be so powerful as to destroy every magpie nest from which cuckoo eggs were ejected. Not only would this be biologically unreasonable, but it would also lead to acceptance behaviour on the part of all magpies and that has not happened (see *Figure* 1). Instead, it is far more reasonable to expect the Mafia to work with less than perfect efficiency, with the result that ejectors, acceptors and abandoners would coexist.

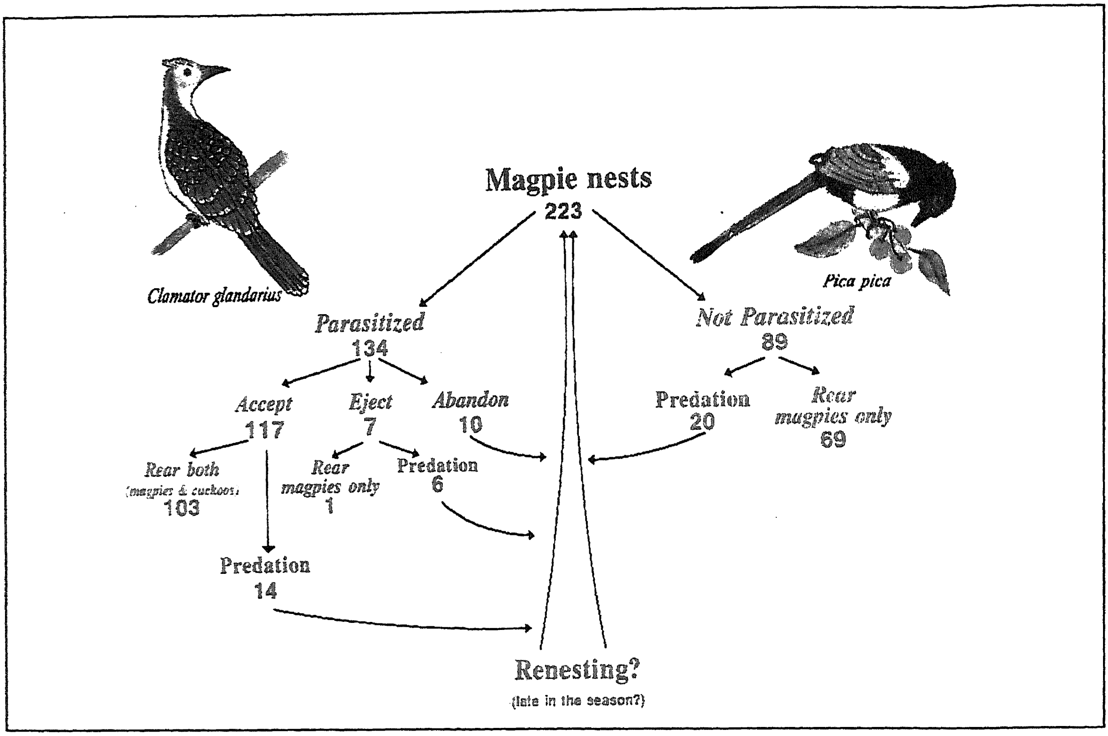


***Figure 1** Rates of parasitization, acceptance, ejection and abandonment in the study population in Hoya de Guadix in Spain during 1991-1992. The host, the black-billed magpie* **Pica pica** *and the parasite, the great-spotted cuckoo* **Clamator glandarius** *are also shown. Data from Soler et al (1995).*

Indeed one can imagine that acceptance begins to pay better if everybody else is ejecting and ejection begins to pay off if everybody else is accepting.

Thus, the prediction of the Mafia hypothesis would not be that acceptors fare better than ejectors but that acceptors should not fare any worse than the ejectors. This latter prediction is supported by both the natural population study as well as the experimental study. That is good evidence for Bird Mafia!

Suggested Reading

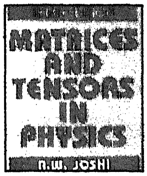
S I Rothstein. ***Ann. Rev. Ecol. Syst.*** **21, 481-508. 1990.**
A H Lotem, H Naklamura and A Zahavi. ***Behav. Ecol.*** **3, 128-132. 1992.**
R Gadagkar. ***Down To Earth.*** **2, 46-47. 1993.**
A Zahavi. ***Am. Nat.*** **113, 157-159. 1979**
M Soler, J J Soler, J G Martinez and A P Moller. ***Evolution*** **49, 770-775. 1995.**
R Gadagkar and M Kolatkar. ***Curr. Sci.*** **170, 115-117. 1996.**

Raghavendra Gadagkar and Milind Kolatkar are at the Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064, India. Raghavendra Gadagkar is also in the Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012, India

A Down to Earth Exposition

Matrices and Tensors Made Easy

Arvind



Matrices and Tensors in Physics
A W Joshi
New Age International Publishers
Ltd, and Wiley Eastern Ltd,
New Delhi, pp.342, Rs.135.

While writing this review, I asked myself: when does one first encounter matrices as a student of physics, and how well equipped is one to handle them? I recalled that the moment of inertia was the first matrix I encountered, and even though I had learnt about matrices in detail at the plus 2 level, I felt rather ill equipped to grasp their significance. Tensors are always projected as difficult and something which mainly general relativists have to learn.

An early introduction to matrix language helps in appreciating concepts like moment of inertia, polarizability, normal modes of systems of coupled oscillators and gives a deeper insight into the unifying principles of classical physics. Understanding the role of rotations in identifying the principal moments of inertia or the basis in which the polarizability tensor (3x 3 matrix) is diagonal, is very important and requires a familiarity with matrices. In the long run, familiarity with the powerful techniques of matrix algebra makes it easy to learn quantum mechanics. Several modern books on quantum mechanics start with systems with a finite number of energy levels and use matrix algebra to introduce quantum concepts instead of following the more standard approach of differential equations. The book by J J Sakurai (*Modern Quantum Mechanics*) is one such example.

Tensors may also be very helpful in handling things as simple as cross products of vectors; the introduction of the completely antisymmetric third rank tensor ε_{jkl} makes life so simple in many elementary calculations involving cross products that one finds oneself using tensors without trepidation! The covariant formulation of electromagnetic theory is a beautiful example where the simple use of tensors gives a lot of insight into the structure of Maxwell's equations.

It is thus very important to have access to a good text book on matrices and tensors which presents the material in a clear and application oriented style. This is the slot that this book fits into; it requires a knowledge of plus 2 level mathematics and introduces most of what is needed to start understanding those areas of physics which use matrix and tensor language. It is best suited for B.Sc. and M.Sc. physics students.

The book starts by introducing elementary notions of vector spaces and defines matrices as acting on these spaces. This in itself is a useful departure (from the physics angle) from the plus 2 level approach of directly dealing with matrices. This makes the transition to higher rank tensors easier. Topics like determinants and the solution of sets of linear equations using matrix techniques, which the

reader is likely to be familiar with at some level, are dealt with in detail. The important and useful procedures of diagonalising symmetric/hermitian matrices using orthogonal/unitary matrices, the significance of eigenvalues and the correspondence of hermitian symmetric matrices with quadratic forms is discussed in fair detail and no background is assumed. Useful ways of partitioning a matrix into submatrices and using this to simplify problems are also discussed.

The subject of tensor algebra and tensor calculus is built up from scratch. Explicit details are provided so that the reader does not get scared. This is important as most students are daunted by tensors. The concepts of covariant and contravariant tensors are developed and applications to special relativity and the covariant formulation of electromagnetic theory are worked out in detail. A preliminary discussion on covariant differentiation, Christoffel symbols and curvature prepares the reader interested in learning more about general relativity.

Most of the problems in the book are rather routine. It may be necessary to have a number of standard problems to reinforce basic ideas but it is equally essential to give thought provoking problems; this is one major shortcoming of the book. Some introduction to the coordinate independent approach may also have been useful. Elasticity could have been discussed while dealing with the applications of tensors.

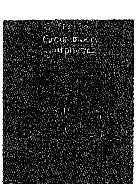
On the whole the book is written with a down to earth approach, giving useful material without getting lost in detailed proofs and at the same time maintaining the requisite mathematical rigour. It should serve as a useful text book/reference for B.Sc as well as M.Sc students. The price is very reasonable.

Arvind, Department of Physics, Indian Institute of Science, Bangalore 560 012.

A Mathematician Looks at Physics

An Introduction to Group Theory and its Applications

Vishwambhar Pati



Group Theory and Physics,
S Sternberg.
Cambridge University Press, 1994
pp. 427. £ 50

That geometry and physics are closely interlinked human endeavours harking back to antiquity is well known. What is not so well known is that group theory, in an implicit way, is also at least as old. For example, the notion of congruence in Euclidean geometry, in modern language would be called equivalence under the group $E(2)$ of rigid motions of the Euclidean plane. Much later, when geometries other than the Euclidean were discovered, they spawned their own

groups of isometries (congruences), such as $PSL(2,\mathrm{R})$ for the hyperbolic plane. At the turn of the last century, Felix Klein's Erlanger program sought to classify and understand geometries as a manifestation of their corresponding symmetry groups. Similarly in physics, for example, the Galilean invariance of classical mechanics, the Lorentz invariance of electromagnetism, or the more recent gauge invariance under various gauge groups have underscored the same leitmotiv, i.e. understand physical laws by their symmetries. This is what the book under review sets out to investigate, starting essentially from scratch.

After a few preliminaries, the author starts off with a purely group theoretic proof of the existence of only five regular (Platonic) solids in three space. It boils down to the classification of all finite subgroups of the three dimensional rotation group $SO(3)$ which do not have a common invariant axis (i.e. are not planar rotation groups, which are also easily classified). The proof involves nothing more complicated than counting a set in two different ways, but already illustrates the power of group theory. The sections 1.9 and 1.10 contain a fascinating excursion into crystallography. Consider the following remarkable fact, which was first empirically observed. The symmetry groups of naturally occurring crystals (which are not Platonic solids in general) did not contain any rotations of orders other than $k = 1,2,3,4,6$. The explanation for this is to be found in section 1.9. Nature has arranged crystal shapes so that their groups of symmetries also preserve three dimensional crystal lattices. This forces only those allowed rotations. Of course, we're still left with the question of what groups do occur as symmetry groups of crystals. Section 1.9 shows that this is basically the classification of finite subgroups of $O(3)$ having only the allowed rotations above, and there are thirty two of them. All but two of these are symmetry groups of crystals occurring in nature, and their table with accompanying pictures occurs at the end of section 1.9. The related question of classifying the crystallographic groups is answered in Appendix A.

If this doesn't strike you as anything more than a curiosity, I would urge you to persist with the representation theory of finite groups in Chapter 2. It is a very clean and efficient account of all the basics, with plenty of illustrative examples and computations of character tables, explicit bases etc. For his or her perseverance, the reader will be amply rewarded in Chapter 3, which contains a detailed discussion of why selection rules occur in physics. The finite group representation theory comes into the study of selection rules for a molecule vibrating in an electromagnetic field, and Raman scattering.

Section 3.8 contains some basics about irreducible representations of semidirect

> Sternberg has done a masterly job of organising the material in such a way that the book is readable from practically any point on.

An imaginative teacher could fashion several interesting courses out of this book at the M.Sc. level.

products via the use of induced representations, and this is immediately deployed to achieve Wigner's classification of the irreducible unitary representations of the universal cover of the Poincaré group. In the physically relevant cases, these turn out to be parametrised by two parameters: m (rest mass) which has any non-negative value, and s (spin), which is allowed to have only half-integer values. The program of Wigner, a kind of Erlanger program for physics, to understand all elementary particles by looking at irreducible representations of the Poincaré group could not explain why only certain masses were found, or how the conserved quantity of charge was to be explained. On the other hand, all of present day physics accepts the validity of this philosophy, and seeks only to find the correct replacement for the Poincaré group. The chapter ends with a group theoretic discussion of parity.

Chapter 4 launches the representation theory of compact groups, which is a very beautiful and complete area of mathematics, and has the flavour of the finite group representation theory discussed in Chapter 2. Some of the proofs (such as the elegant proof of the existence of Haar measure using the Mean Ergodic theorem, or of the Peter-Weyl package for the regular representation) are relegated to the Appendices, and quite wisely so. This is because the immediate application of $SO(3)$ representation theory to the explanation of the various quantum numbers of orbital electrons in an atom is a very beautiful and intellectually satisfying exercise which begins in Section 4.5. In fact, if you were baffled by the numbers $2(2l+1)$ of electrons in a given L-subshell while studying school chemistry, the answer lies in the representation theory of $SO(3)$, as you will discover in this section. Then the story moves on to the Clebsch-Gordan decomposition of two irreducible $SU(2)$ representations, and its physical implications, e.g. isospin.

Chapter 5 moves to the classification of the irreducible finite dimensional representations of $SU(n)$, $SL(n,C)$ and $GL(n,C)$. This is done via the decomposition of the rth tensor power of C^n into its irreducible components under the natural representation of the symmetric group S_r. The remarkable application of the adjoint representation of $SU(3)$ to the 'Eight-Fold Way' of Gell–mann and Neeman and its consequent prediction of the Ω particle is discussed in 5.10. Section 5.13 goes into gauge theory, but not in too much detail.

Sternberg has done a masterly job of organising the material so that the book is readable from practically any point on. The proofs are always the slickest possible, and there are ample examples along the way so that a general reader does not lose himself or herself in abstraction. (Note the beautiful introduction of a vector bundle via molecular vibrations in 3.2, or the neat way of explicitly constructing

> It is a very clean and efficient account of all the basics, with plenty of illustrative examples.

a Haar measure on a linear Lie group in 4.1.) The only prerequisites are a sound understanding of undergraduate analysis and linear algebra. While this book could be considered too sophisticated for the undergraduate level in an Indian university, an imaginative teacher could fashion several interesting courses out of this book at the M.Sc. level.

On the downside, the index is woefully inadequate for a book of such encyclopaedic reach. In fact, since Sternberg eschews the usual definition-theorem-corollary format of mathematics books, the index needs to be exhaustive. For example, *Poincaré Group, spin and angular momentum* do not exist as entries in the index! Another inadequacy is the absence of exercises, which could have been used to provide the sometimes very elementary demonstrations that Sternberg goes through in painstaking detail. There are also more typos than one would expect from such a reputed publisher.

Finally, I hope the author will write a sequel to this book where the Cartan-Weyl theory and classification of groups such as Spin and Spin_c and their representations, which are of great interest to physicists, and finally gauge theory are dealt with in greater detail. Meanwhile, happy reading, or for that matter, even browsing!

Vishwambhar Pati is with the Indian Statistical Institute, Bangalore 560059.

The Strength of Materials

Without Stress or Strain

Gangan Prathap

The New Science of Strong Materials or Why you don't fall through the floor
J E Gordon
Reprinted in Penguin Books, 1991.
pp.287, Rs.195.

Stress and strain are words used to describe the mental condition of human beings. In this sense, the words are used interchangeably to mean the same thing. In science and engineering, these words are given distinct meanings and the entire science underlying structural engineering rests on these basic distinctions. The structural engineer's craft rests on two pillars: form (or how shape and size and manner of arrangement of structural material provides efficient design) and substance (the nature of material(s) out of which the structure is fashioned). An earlier book by Gordon, 'Structures or Why Things Don't Fall Down' (also available in Penguin edition), dealt with the former aspect; the book reviewed here deals with the latter issue.

"Why do things break?" is the substantial question addressed in this book. "How is this

> "Why do things break?" is the substantial question addressed in this book. "How is this to be interpreted in terms of the physics and chemistry of the material constitution?"

to be interpreted in terms of the physics and chemistry of the material constitution?" Gordon surveys the field from its historical origins to the most recent advances in materials design (yes, engineering new materials for optimal structural performance) using as little mathematical equipment as possible, meeting Lord Ashby's criterion of "technological humanism"- "the habit of apprehending a technology in its completeness."

Chapter one asks awkward questions on the new science of strong materials, which has taken shape only in this half of the century. It sets the stage for what follows.

The remaining ten chapters are arranged into three parts. Part one deals with elasticity theory and the theory of strength, where the engineering definitions of stress and strain (from Galileo to Hooke and then to Young) are clearly spelled out and the manner in which an understanding of the distributions of stress and strain in a structure assists in predicting its structural performance. How strong any given material is and what the causes of weakness are, come next; a fair amount of attention is devoted to the role that cracks and dislocations play and how their growth under adverse conditions of stress concentrations can certainly lead to distress.

Materials scientists divide themselves into two factions, the metallurgists and the non-mets, not entirely unexpected because the properties of metals and non-metals are evidently so different. This division has "run right through the history of technology" and Gordon very fairly divides the remaining and greater part of the book into these two traditions. Part two discusses how non-metallic materials can be toughened to stop cracks. It deals with such common place materials as timber, cellulose and plywood and proceeds from here to the more recent wonder materials — the world of composites. Part three is an intimate study of iron and steel, which made the industrial revolution possible, and of metals in general. The book ends with some "technological forecasting", and happily it turns out that chapter 11, "The materials of the future" was supported by a sub-title which Gordon could change from a defensive "or how to guess wrongly" (first edition of the book in 1968) to an optimistic "or how to have second thoughts" when he revised the book in 1976. Most of his prophecies became accomplished facts. "Getting to the moon... was a very expensive

> Gordon surveys the field from its historical origins to the most recent advances using as little mathematical equipment as possible, meeting Lord Ashby's criterion of "technological humanism".

> "Getting to the moon... was a very expensive way of developing non-stick frying pans", sums up the spirit of "technological humanism" that Gordon stamps on this account of materials science.

way of developing non-stick frying pans", sums up the spirit of "technological humanism" that Gordon stamps on this account of materials science.

The end of the book carries two appendices on various kinds of solids and on a simple beam formula. There is a list of books suggested for further study, much in the same way *Resonance* instructs its contributors to end their essays (as this review also dutifully does, borrowing from the same list). A helpful index is also provided.

I found and relished Gordon's two books much after completing my formal degree in engineering. I couldn't help feel a sense of regret and incompleteness that my engineering education never did of itself provide a fabric into which technology could be woven so that I could claim to have been given a liberal education (to paraphrase the long quote from Lord Ashby's 'Technology and the Academics' which Gordon places as a motto at the head of his book). I hope that readers of *Resonance* will get hold of these two books, even if they are not structural engineers or materials scientists, to educate themselves of the "sometimes curious and entertaining ways in which scientists isolate and solve problems."

Suggested Reading

S P Timoshenko. History of the Strength of Materials. McGraw-Hill, 1953.

History of Technology. Oxford University Press. 1954.

D'Arcy Thompson. On Growth and Form (abridged edition). Cambridge University Press. 1961.

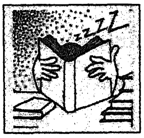
E Torroja. Philosophy of Structures (translated from the Spanish). University of California Press, Berkeley, 1962.

B R Schlenken. Introduction to Materials Science, John Wiley, 1974.

Gangan Prathap is with National Aerospace Laboratories, Bangalore 560 017 and Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064.

"Agata Mendel ... jokes that most contemporary molecular and developmental biologists reason like a child who, because turning the knob on the television set makes the picture appear, concludes that the knob 'causes' or 'programs' the picture, and then goes the next absurd, step of trying to understand the mechanism of television by chemically analysing the knob" from *Theory and practice of genetic reductionisim –from Mendel's laws to genetic engineering,* an essay by R Hubbard in ' *Towards a Liberatory biology* edited by S Rose.

Books Received

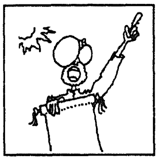


Tradition, Science and Society
S Balachandra Rao
Navakarnataka
1990, Rs.15.

A Textbook of Biology - Volume 2
P K G Nair, K P Achar, M J Hegde and S G Prabhu
Himalaya Publishing House
1996, Rs.150.

Indus Script - Its Nature and Structure
B V Subbarayappa
New Era Publications.
1996, Rs.230.

Information and Announcements



The 37th International Mathematical Olympiad

As has already been announced in these pages earlier (*Resonance*, Vol.1, No.1, 1996), India is hosting the 37th IMO from 5 July to 17 July, 1996 in Mumbai. (The venue was shifted from New Delhi to Mumbai in March.) About 75 countries are expected to participate.

The problems posed in the final paper are selected through an interesting procedure. Each participating country (except the host) submits upto six problems to the *problem short-listing committee* of the host country. The committee receives about 120-150 problems from which it prepares a short-list of about 30 problems taking care to see that all the areas — algebra, combinatorics, geometry and number theory — are represented. The most important single criterion for a problem to be short-listed is that it should be new and not have been published anywhere. This is a very hard thing to ensure.

The jury which consists of leaders of all the participating countries arrives three days

ahead of the contestants and the deputy leader to select the final six problems that go into the IMO paper from the short-list. Thus each country participates in the making of the IMO paper.

The opening ceremony is on 9 July and the contests will be held on July 10-11. The closing ceremony where the medals are awarded will take place on 16 July.

C S Yogananda, NBHM (DAE), Department of Mathematics, Indian Institute of Science, Bangalore 560 012, India.

Indian Statistical Institute

It began as a small room in Presidency College, Calcutta in 1931 and is now one of India's major academic institutions. The Indian Statistical Institute was founded by Prasanta Chandra Mahalanobis, a physicist, who realised the importance of statistics when he was completing his Tripos in Cambridge, England in 1915. He established a statistical laboratory in the Physics Department of Presidency College, Calcutta in the 1920's; this laboratory later became the Indian Statistical Institute, registered as a learned society in 1932. From then on, it grew in size and stature and today it is one of the leading institutions of its kind. It has a strength of over 250 faculty members and over 1,000 supporting staff. It is a major presence in Calcutta, Delhi, Bangalore and other cities, as a research and educational institution and as a project and consultancy centre. It was declared an institution of national importance by an act of Parliament in 1959 and was vested with powers to award degrees and diplomas in statistics. Recently, by an amendment of this act, it was also vested with powers to award degrees and diplomas in mathematics, computer science and quantitative economics and such fields related to statistics.

Early in his career, Mahalanobis realised the key role statistics can play in scientific investigations and national development. He realised from his own work, and from developments elsewhere in the world, that the then young subject of statistics was a potentially fertile area for research in theory, methodology and applications. He was of the opinion that relevant statistical theory and methodology can develop only in an environment where statisticians were both aware of its applications and participated in them. Thus he believed that in addition to researchers trained in statistical theory a statistical institute must have natural and social scientists engaged in quantitative research in their own areas. This would help the natural and social scientists get statistical expertise to design their studies and analyse their data, and the statisticians to understand and work on theoretical and methodological problems of a genuine nature. With this point of view, he established a number of science units in the institute.

The headquarters of the Institute are in Calcutta, with centres in Delhi and Bangalore. A variety of research, training and project activities take place in Calcutta and the other centres. The institute has units in the cities of Baroda, Bombay, Coimbatore, Hyderabad, Madras, Pune and Thiruvananthapuram, which offer services to the industry in quality control, operations research and management.

The Institute has been offering formal courses in statistics and related fields leading to certificates and diplomas since the 1930's. Since 1960, the Institute has also been offering courses leading to B.Stat., M.Stat. and Ph.D. degrees. M.Tech. courses in computer science and quality, reliability and operations research have been introduced subsequently. An M.S. programme in quantitative economics is being introduced in 1996. Admissions to these courses are based on academic records and performance in All-India selection tests and interviews. The selection tests are generally conducted in May and advertisements for these appear in late January in major national newspapers. Only very meritorious candidates get admission into these courses and all of them carry stipends. The B.Stat. programme is offered only at the Calcutta centre. An idea of the level of the selection tests for B.Stat. can be obtained by consulting the booklet prepared by the Indian Statistical Institute titled "Test of Mathematics at the 10+2 level" (New Delhi: Affiliated East-West Press Ltd., Price Rs.49). The M.Stat. programme is offered at the Bangalore, Calcutta and Delhi centres. The Ph.D. programmes in economics, mathematics and statistics are offered at Bangalore and Delhi and the Ph.D. programmes in computer science, economics, mathematics and statistics are offered at Calcutta.

In the statistics degree programmes run by the Institute, the mathematical and theoretical bases of statistics are established. In addition, the practical relevance of the methods are also emphasized by means of practical work, projects and statistical analysis of live data, from a variety of real-life situations. The students are also trained in modern computer work. At the end of their courses in the Institute, students are equipped to join an academic or scientific research programme or in service in the public or private sector dealing with applications of statistics, mathematics, economics and computer science in a variety of fields.

T Krishnan, Indian Statistical Institute, 203, BT Road, Calcutta 700 035, India.

Haldane's view ... J B S Haldane was once asked what the study of biology could tell one about the Almighty. "I'm really not sure," said Haldane, "except that He must be inordinately fond of beetles." There are thought to be at least 300,000 species of beetles. By contrast, there are only about 10,000 species of birds. (From *Lessons from Biology* by Francis Crick. *Natural History* 11/88).

Acknowledgements

Resonance gratefully acknowledges the help received from the following individuals: V Pati, S Ramasubramanian and Jayant Rao.

Errata

Resonance, Vol. 1, No. 4 (1996)

Page 36: Optical illusions: (left) Each crosshatched line is a straight line though it appears distorted; (right) a system of concentric circles that appear to be ovals.

Page 64: The question on 'Self-Copying Program' in the *Think It Over* section was posed by S K Ghoshal, not V Rajaraman

Page 66: Optical illusion: The optical illusion is a system of concentric circles and not a spiral pattern.

Annual Subscription Rates

	Personal	Institutional
India	Rs 100	Rs 200
Third World Countries	US $ 25	US $ 50
Other Countries	US $ 50	US $ 100

Send your subscriptions by DD or MO in favour of "Indian Academy of Sciences" to Circulation Department, Indian Academy of Sciences, C V Raman Avenue, PB No. 8005, Bangalore 560 080, India.

Edited and published by V K Gaur for the Indian Academy of Sciences, Bangalore 560 080. Printed at Thomson Press (I) Ltd., Faridabad 121 007.

John (Janos) von Neumann was born in Budapest, Hungary, on December 8, 1903. His talents were discovered early by his school teachers and he was privately tutored in mathematics by university teachers. His first paper appeared before he was 18 years old. Interestingly, he obtained his doctorate in Mathematics in Budapest about the same time as his degree in chemical engineering in Zurich!

He was a Privat Dozent *for three years at the Universities of Berlin and Hamburg before migrating to the USA as a visiting lecturer at Princeton University in 1930 where he soon became a professor. When the Institute for Advanced Study was established in Princeton in 1933 he moved there as a permanent professor. He played a leading role in the Los Alamos atomic bomb project from its beginning.*

Von Neumann's brilliant and original contributions cover the wide spectrum of scientific thought of his times. He combined in his mind several abilities that one rarely finds in one intellect — a feeling for the set-theoretical, formally algebraic basis of mathematical thought, the knowledge and understanding of the substance of classical mathematics in analysis and geometry, and an acute perception of the potential of modern mathematics for the formulation and solution of problems in other branches of human endeavour. This last mentioned ability is amply demonstrated in his most original creation, Game Theory.

After his early work in the foundations of mathematics wherein he formulated the von Neumann-Bernays-Gödel set theory and gave a new definition of ordinal numbers he turned his attention to providing a rigorous mathematical formulation of quantum mechanics. His book, Die Mathematische Grundlagen der Quantum Mechanik *is still a classic on the subject. This work motivated him to study operators on Hilbert spaces where he initiated the study of certain rings of operators, von Neumann algebras. This is now a very active area of research and has had profound impact on the development of various branches of mathematics.*

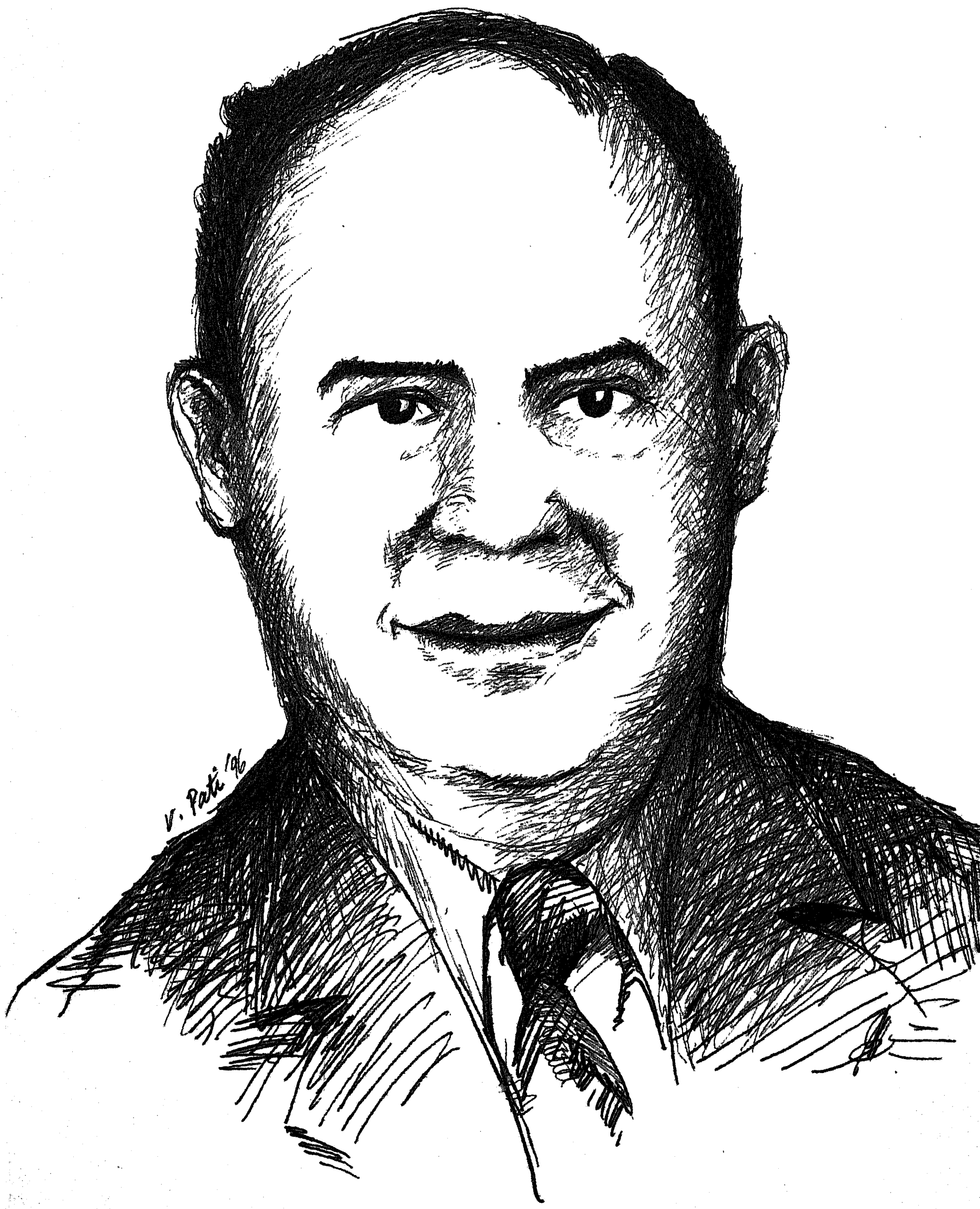
His other major contributions to mathematics include the solution to Hilbert's fifth problem for the case of compact groups and the first proof of the mean ergodic theorem. In Game Theory he formulated and proved the minimax (minimising the maximum losses) principle. He did pioneering work in devising electronic computing machines.

He was famed for his remarkable ability to solve in his head, problems that made other mathematicians turn to pen and paper or even calculators.

Von Neumann stories abound. After an automobile crash in Princeton he came out of his wrecked car to explain, "the trees on the right were passing me in orderly fashion at 60 miles per hour. Suddenly one of them stepped out into my path!"

Von Neumann and his first wife, Marietta Kovesi had a daughter, Marina. His second wife, Klara Dan later became one of the first coders of mathematical problems for electronic computing machines.

He died in Washington on February 8, 1957.



John von Neumann
1903-1957